AD-A119 767

NAVAL AIR DEVELOPMENT CENTER WARMINSTER PA AIRCRAFT --ETC F/G 17/2
VOCAL AND MANUAL RESPONSE MODES: COMPARISON USING A TIME-SHARIN--ETC(U)
JAN 79   C P SKRIVER

UNCLASSIFIED NADC-79127-60

NL
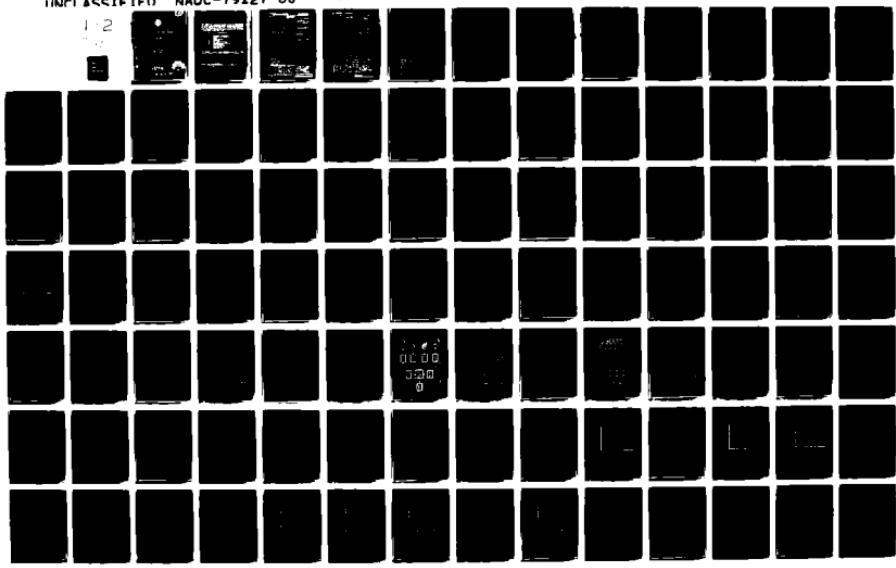
VOCAL AND MANUAL RESPONSE MODE COMPARISON USING A TIME-SHARING PARADIGM

Christine P. Skelton
Aircraft and Crew Systems Technology Department
NAVAL AIR DEVELOPMENT CENTER
Warminster, Pennsylvania 18974

JANUARY 1979

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

1. REPORT NUMBER
NADC-79127-60

4. TITLE (and Subtitle)
VOCAL AND MANUAL RESPONSE MODES:
COMPARISON USING A TIME-SHARING
PARADIGM

5. TYPE OF REPORT & PERIOD COVERED
FINAL

7. AUTHOR(s)
CHRISTIAN P. SKRIVER

9. PERFORMING ORGANIZATION NAME AND ADDRESS
HUMAN FACTORS ENGINEERING DIVISION
AIRCRAFT AND CREW SYSTEMS TECH. DIR.
WARMINSTER, PA 18974

10. Refer to Block 18

11. CONTROLLING OFFICE NAME AND ADDRESS
NAVAL AIR SYSTEMS COMMAND
WASHINGTON D.C. 20361

12. REPORT DATE
JAN 1979

13. NUMBER OF PAGES
132

14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)
HUMAN FACTORS ENGINEERING DIVISION
AIRCRAFT AND CREW SYSTEMS TECH. DIR.
WARMINSTER, PA 18974

15. SECURITY CLASS (of this report)
UNCLASSIFIED

16. DISTRIBUTION STATEMENT (of this Report)

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

62757N, F57525, 63701N, WO542-PN
WF57525492, ZH302, WO542001.06, RD505

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)
AUTOMATED SPEECH TECHNOLOGY
AUTOMATIC SPEECH RECOGNITION
CREW STATION DESIGN
HUMAN FACTORS

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)
In the increasingly complex area of man-machine interface,
the human operator is often required to perform several
tasks which demand full attention simultaneously, thus
overloading the operator's hands and eyes.
The recent commercial availability of speech recognition
devices has opened up the possibility that a vocal
response could act as an extra pair of hands for the operator.
Proponents of automated

20.

that vocal response would be independent of other channels; would interfere less with other activities and would be faster than manual response mode. The present study was designed to test the effectiveness of a voice response as compared with a manual response in a time-sharing environment.

Seventeen subjects participated in the study. All subjects performed a digit entry task concurrently with a tracking task. Subjects also performed all three tasks-- vocal entry, manual entry and tracking--in isolation. The digit entry task required the subjects to vocalize a digit or press a key corresponding to a digit presented on a CRT display. Vocal input was recorded via a voice recognition device. Manual input was through a sixteen key keyboard. The compensatory tracking task required appropriate left- right movements of a handcontrol to maintain the position of a cursor on the center of a horizontal track which was also displayed on the CRT. In addition, a feedback indicator presented the subjects with knowledge of results. Following a training period, the subjects performed a series of fifteen three minute trials in which their performance was tested on the three tasks in isolation. In addition, the manual task and the vocal task were performed concurrently with the tracking task.

A randomized block factorial design was used in evaluating the effects of three levels of response alternatives (four, eight, or sixteen digits in a set) and three levels of tracking task load (tracking alone, tracking with vocal entry, or tracking with manual entry) upon tracking performance. A randomized block factorial design was also used in evaluating the effects of three levels of response alternatives , two levels of response mode and two levels of task load upon the performance of a digit entry task. The dependent variables recorded in this study were: 1) digit entry response latency interval measured from start of digit presentation to the registration of a response to that digit; 2) digit entry errors, and 3) tracking root mean square error.

The results indicated that manual entry performance declined with increased number of response alternatives, and that both manual and vocal performances deteriorated with increased task load, although deterioration was significantly less with the vocal input mode. Tracking performance also decreased with dual-task load. Response alternative loading had no significant effect on the performance of the vocal entry or tracking tasks. The manual mode was found to have a higher error rate than the vocal mode.

Results from this study demonstrate that the vocal response mode is dramatically more effective in terms of speed and accuracy than manual mode when it comes to high task difficulty in a time-sharing situation. The hypothesis that the vocal response mode is faster than the manual response mode was rejected in favor of a more general alternative hypothesis that speed of response mode is a function of task load and the number of response alternatives.

ACKNOWLEDGMENTS

Accession For

NTIS GRA&I	X
DTIC TAB	☐
Unannounced	☐
Justification	

By
Distribution/
Availability Codes

Dist	Avail and/or Special
A	

ii

TABLE OF CONTENTS

iv

LIST OF TABLES

v

LIST OF ILLUSTRATIONS

CHAPTER I

INTRODUCTION

The concept of attention has historically been of concern not only to experimental psychology in general but also to a number of applied fields which must deal with the practical concerns of human operation in highly automated systems. Flying high performance aircraft, in military, commercial, and private aviation is one such applied area. Aviation psychology--a special branch of applied psychology--attempts to understand and resolve problems which arise in the increasingly complex operations required by pilots in this area. One common problem is the requirement that the human operator in aviation perform two or more tasks which demand his attention simultaneously.

The problem of divided attention, of course, goes beyond aviation to any modern highly technological system with sophisticated instrumentation. High performance aircraft operation, however, which has imposed extremely complex performance requirements on system operators, offers a representative environment in which to study this problem. Chapanis, shifting among various instrument indicators, switches, and controls on the various planes, stated that instrumental maze and design

1

accounted for the large amount of "human element" accidents. Since then, the number of instruments in fighter aircraft have further multiplied, and the complexity in avionics has reached the point at which the demands on operator attention frequently surpass the operator's limits of capability. On-board computers have in many instances added to the instrument confusion by overloading operators with more complex chores than they can effectively deal with. Lane, Strieb and Wherry, Jr. (Note 1, p. 1) report that

> within some highly automated crew-stations, crew-members are able to exercise only a fraction of the repertoire of functions provided by on-board computers; at the same time they are often overwhelmed by having to perform numerous other functions that are frequently more adaptable to automation than those already computerized.

Yet, these functions may be resistant to computerization because the task can not be or has not been defined and thus are left to the attention of the human operator.

One of the most critical problems in an advanced airborne weapons system is that human operators are still required to initiate missile launches. While many aspects of this task are computer augmented, the operator must

attend to the most critical aspects of target acquisition, weapons selection and launch in order to insure mission success and crew safety. The timing and accuracy of the operator's performance, while concurrently attending to tracking, arms selection and launch, are very important for the success of mission goals. These tasks depend on the operator's manual and visual performance while leaving his vocal capability relatively unloaded. If the voice could be utilized as well, the total performance of the operator might improve considerably.

A recent development in computerization which appears relevant to both the alleviation of activity overload and the adaptation to automation of functions not yet computerized, is the commercial availability of voice recognition and synthesis equipment.

Although man-robot communications have long been a favorite object of science fiction, it is not more than a decade since speech communications experts predicted the coming of serious application of speech in direct man-machine interactions. In a panel discussion reported by Lindgren (1967), M.R. Schroeder predicted that voice-actuated machines and even machine-actuated people would soon be a reality, but warned that research into speech oriented technology, which up to then had been

carried out mostly by physical scientists, must shift its emphasis to investigations of basic human behavior, human capabilities and their limitations.

Besides the obvious fact that speech is one of man's most natural modes of communication, there are other reasons to pursue further investigations of automatic speech technology. Hill (1972) listed additional advantages of speech communications with machines: less interference with other activities than other channels; providing an extra channel for multi-modal communications; using less panel space; being compatible with available communication systems; being independent of factors that normally affect sight and reach, etc. It should be noted that some of the above claims have not been fully substantiated yet by any experimental studies and that conclusions must be deferred until controlled investigations can be made.

In a recent review of automatic speech technology Feuge (Note 2) stated that the solution to crew-station overload problems is a new mode of inputting and outputting data into a system that "frees the hands and eyes for other tasks; does not disrupt attention; is faster than present modes; and is equally (or more) accurate than present modes" (p. 2). Feuge suggests that the automatic speech recognition devices meet these criteria, although no formal tests

of these hypotheses were presented in his review.

Other ongoing investigations are probing the possible practical uses of various voice recognition and synthesis systems that are presently being used on a limited scale only. According to Beek, Neuberg, and Hodge (1977)

> These applications are limited to a small vocabulary, speaker-dependent, isolated word recognition system. These highly reliable systems allow for hands-free source data entry of the digits and a limited set of control words. As such, the voice data entry system eliminates manual transcription and keying operation (p. 2).

Beek, et al (1977) reported that a number of military applications can be expected in the near future: in communications systems; for automatic speaker verification; as an integral part of military training; word recognition for military tactical data handling; voice recognition and synthesis for aircraft cockpits, etc.

It appears then that for certain operator functions the addition of a "vocal channel" could increase an operator's capacity, reduce workload and possibly reduce the amount of instrumentation clutter. Vocal input/output devices should therefore be taken into serious consideration as alternative

means of alleviating crewstation work overload.

At this time, however, since most of the above systems are still being evaluated, it must be emphasized that very little performance data are available for consideration and that available studies are either incomplete, inconclusive or contradictory. Especially notable is the fact that the studies that will be reviewed in the following pages deal mostly with auditory stimulus input or with verbal and manual responses to visual or auditory stimuli, while _vocal_ response output implies in this study that the subject is using the voice as an alternative response mode, "an extra pair of hands." It is common to find that in the study of the locus of interference in information processing, the encoding and identification stages have been studied, more so than the selection of response or the actual execution of response. Only a few studies were found in the literature that reported the effectiveness of the vocal response mode in comparison with the manual response mode in a dual-task performance situation.

If the new voice recognition and synthesis systems are going to be considered for inclusion in the cockpit, for instance, it is necessary to find out if a voice response is more efficient than a keypress response when the operator's hands are already engaged in the primary task of tracking.

The present study was undertaken in order to investigate how a voice mode interacts with performance in a time-sharing situation, allowing a comparison of the task load associated with manual versus vocal controls. Such a study not only has theoretical implications, but would also be most useful in establishing guidelines for design of systems incorporating automated speech recognition devices.

Review of the Literature

As stated earlier, relatively few studies have investigated vocal response as an alternative response mode, but since the critical variables in the present study are: response mode, response alternatives, and time-sharing, there are studies that deal with these variables, which have proven useful and applicable to this study.

Time-sharing and Task Interference

The interaction of two simultaneous performances was investigated as early as 1887 when Paulhan demonstrated that he could recite a familiar poem aloud and at the same time write another poem or do multiplications on paper without losing track of one or the ther (Woodworth and Schlosberg, 1954). Another early attempt to deal with the efficiency of a dual performance is reported by the same authors (p. 88), in which Sterzinger in 1928 conducted an experiment in which

he had subjects listen to a story being read aloud while they were attempting to add a single column of numbers. The subjects were then told to stop adding and to write down the story. The results were calculated according to Sterzinger's index of efficiency, which showed a range from 30 to 90 with a group average of 60 on the efficiency index. The same authors attempted to define divided attention:

> . . . doing two things at once. The subject attempts to perform two tasks simultaneously. The double task-set may be represented by O_1 and O_2 and the formula written thus, R_1 R_2 = f $(O_1 O_2, S_1, S_2)$ and the question is whether R_1 R_2 occurs at all and if so how efficently (p. 73)

Models of Time-sharing Performance

"The concept of selective attention is as important in learning to fly high-performance aircraft as it is to current experimental psychology" (Gopher and Kahnemann, 1971, p. 1335).

Although the study of divided attention can be traced back to the infancy of experimental research, it was not until the 1950s and 60s that significant efforts were made by experimental psychologists to relate investigations of

attention to the relatively new theories in communications and adopt many of the methods from this engineering field. Information processing models have borrowed extensively from communications theory, metrics and terminology in that the concept of communication channels, bandwidth, senders, receivers, etc. could easily be applied to the human system as well as the telephone system. Most theorists have used the flow chart to illustrate their particular aspect of information processing, such as visual perception, memory storage, information metrics, problem solving, linguistics, to name a few (Alluisi, 1970; Haber, 1974).

In the field of perception and information processing, especially investigations of dichotic listening, Broadbent (1958) is generally acknowledged as the prime mover. His "filter theory" proposes a mechanism of selective attention that acts as a filter, allowing some information to filter through but excluding or short-term storing the remainder in order to protect information channels against overload, thus admitting information input of only one channel at a time but allowing switching between channels. This "single-channel" model--an engineering construct that implies that only one "thing" at a time can be processed or attended to--has been investigated by experimenters, and has been supported, challenged and modified by several theorists.

In order to account for the admittance of irrelevant, unwanted messages, which the filter in certain situations does not screen out, Treisman (1960) modified the filter theory, suggesting a perceptual filter where some information messages are weakened but not totally eradicated, merely attenuated. Treisman also added a "dictionary unit" mechanism, where each signal that comes through the filter unattenuated, is analyzed, matched and assessed for importance. Still, only the fully assessed signal will be dealt with.

Broadbent (1971) subsequently revised and expanded his model along the "attenuation theory" lines, but he also noted that psychological models and hypotheses have become so complex that it is nearly impossible for the experimental researcher to test the different versions against one another.

Deutsch and Deutsch (1963) challenged the single-channel bottleneck and proposed that all stimulus input is analyzed, but that a message must pass a certain threshold value in order for the arousal mechanism to be activated. In case of several incoming above-threshold stimuli, a selection process takes place and the message of the most important weight gets preference and is then attended.

Norman (1968, 1969) incorporated a similar re-

sponse-selection theory in his model, which assumes central memory units where inputs are matched and processed both on a physiological level and on a pertinence level. The combined value of these memory data determines what message will be attended.

Neisser's (1967) model which processes information through analysis-by-synthesis, contains a central analyzer where the input is held briefly in a very-short-term memory. If over-learned, the response is immediate; if not, further processing is needed in a focal attention unit and matched in long-term memory.

Welford's (1971) is still a single-channel model which provides for an arousal mechanism located in the brain, allowing the neurons to become increasingly more responsive to stimulus input. But continuous arousal is detrimental to performance: external signals become diffuse and this lowers the channel's capacity to handle the message, "leading to slips of the tongue and other minor confusions of everyday life" (p. 21).

Kahneman's (1973) model which is essentially a "capacity theory", incorporates an arousal mechanism with an "allocation policy" (p. 150-153) to allocate available processing capacity to ongoing activities, suggesting that parallel processing of concurrent messages is possible in the

perceptual system and that the amount of capacity is not a constant "but perception draws on a common pool of capacity, and the ability to carry out detailed analysis of several units is limited" (p. 129).

Teichner (1968, 1974) posits an expanded version of the perceptual filtering model based on physiological observations. Stimuli from both the external and internal world are fed into a sensory processor for recoding and then transferred to short-term memory, where some information will decay, while the remainder is transferred to the attention mechanism. This is a model of a system with a variable capacity rather than a limited capacity channel. Teichner (1974) notes that

> At a given level of human system analysis, the only differences are in the degree of loading or activity of the subtasks. A theory of human performance, therefore, must be a theory of subtask functions and relationships. And the level of human system analysis required must be that which will permit predictions of the performance of the human system at some desired level of accuracy (p. 2).

Norman and Bobrow (1975, 1976) in two recent

theoretical papers proposed a method for analyzing the trade-offs in performance that can occur when a person attempts to perform several activities at the same time. This is a limited capacity model in which the operator must draw upon a limited resource for processing information. Processes may be either data-limited or resource-limited. They argue that conclusions regarding interference among component processes must be made with caution. If a process is data-limited, then allocating any more or less capacity to it, for example, will not change performance and any inferences regarding the interaction of a dual task is suspect. If a task is too easy or too hard, then any comparison with a second task would be difficult to interpret. Within the resource-limited area for each task, the effects of trading off effort among tasks can be evaluated. Interference in this model is due to competition for specialized processing resources, such as memory or processing effort, and any conclusions require a detailed analysis of task structures.

Navon and Gopher (Note 3) have proposed a theory of human performance based on the idea of a multiple resource information processing model. Each resource has limits in its capacity as to how much it can process and to what type of task components it may process. The resource may also at

any moment be allocated among several processes. The authors propose that their processing mechanism may operate in parallel as well as in sequential order. They argue that research which supports serial stage models as well as single-channel models, can be accounted for by the Navon and Gopher approach. This type of model does suggest that efforts to find single indices of channel capacity, work-load, or a standard reference task, may be hopeless. The most fruitful research objectives would be

> to find out what types of resources exist and to
> map out demand compositions of different tasks by
> pairing each of them with each of the other ones.
> If the range of tasks used is diverse enough,
> every type of resources will presumably be
> competed for by at least two tasks. Interference
> data from all the task combinations may be
> sufficient then for recovering by means of
> multidimensional scaling a resource space as well
> as the demand vectors associated with each of the
> tasks (p. 47).

It is almost inevitable that a brief summary of such a vast and ambiguous topic as theories of attention and information processing ends with a plea for more research

definition of the concept of attention. It seems likely that the word "attention" has a different connotation for the experimenter in dual-task problems, in vigilance, in physiological methods, etc. Is there an actual and measurable rather than a semantic difference between "The selection of messages for synthesis". . . and "the operation of a filter"? (Kahneman, 1973,p. 126).

Research into Time-shared Performance

The problem of determining how two relevant variables interact or interfere with one another and affect performance has been the subject of much investigation, often with seemingly contradictory findings.

In a series of experiments Mowbray (1952, 1953, 1954) demonstrated the failure of "doing two things at once." In the first study the subjects were presented with one visual stimulus in the form of sequences of letters or numbers concurrent with an aural presentation. The subjects were supposed to detect missing items in the sequences. In the second experiment Mowbray used a simultaneous visual and auditory prose reading of three levels of difficulty, followed by a comprehension and retention test. The subjects in the third study were presented with problems where the solution had to be given directly or indirectly in

the instruction both visually and aurally in pairs. Mowbray found the resulting performances greatly impaired. The auditory errors were consistently higher than for the visual tasks; the letters generally fared less well than the numbers and the poorly understood message in a pair was seemingly ignored. When two easy readings were paired, both could be comprehended, showing that parallel processing is possible if the task is simple enough. However, when the simple task was paired with a difficult one, it was the more complex task that received the attention, which would agree with the single-channel concept: a difficult task would occupy the single-channel completely and tend to exclude or attenuate messages of easier caliber.

Schvaneveldt (1969) reported a couple of experiments involving manual and verbal responses in a dual task performance using two similar tasks and three levels of information manipulation (0, 1, or 2 bits). In Experiment II the manual task required the subject to respond to the onset of one of four displays with a press of one of four buttons arranged in a corresponding fashion. In the 0-choice condition, the same display was turned on, while in the 4-choice task any one of the four could come on. The vocal task required that the subject read the numeral, which was also the onset, that was presented on the display, and to

transform the number into a corresponding letter: A for 1, B for 2, C for 3, and D for 4. The 0-choice condition used the A-1 pair; the 2-choice task used the A-1, B-2 combination; and the 4-choice used the entire set. Each of the three levels of stimulus alternatives was combined in a factorial design. The subject was required to press a button with the corresponding finger and to call out the correct letter.

The results of the experiment indicated that the average reaction time for both manual and vocal increased linearly with increased information load ($\log_2$ of number of choices). The manual conditions were also faster in reaction time than all of the vocal conditions. Since the vocal task had an additional numeral-to-letter transformation component, the validity of the manual-vocal comparison can be questioned. The data did not seem to support nor disprove either the single-channel or the parallel processing hypotheses.

In Kahneman's (1973) analysis of the above experiment, he made the observation that although the experimenter had defined the pairing of the tasks as two distinctly different response units, the subjects may have responded as if one combined task. The evidence Kahneman presented indicated that the responses were grouped with the vocal trailing the

manual response. The vocal RT became increasingly delayed beyond the manual RT with higher levels of choice.

Kahneman suggests that although interference always occurs in dual-task performance, the subject can "protect" one of the tasks so that it is performed almost to the same level as it would be if performed as a single task. Thus, the interference effect would only affect the subsidiary task. This suggests that the two tasks are performed in parallel rather than in single-channel organization. Parallel processing, according to Kahneman, can be switched to a "sequential strategy" in order to prevent failure by overloading the system and that "jamming of the system is not permitted to occur. . . . The choice of processing mode depends at least in part on the load imposed by the competing activities" (p. 201).

In an investigation of attention sharing in dual-task performance Cliff (1971, 1973) selected a visual and manual compensatory tracking task coupled with an auditory/verbal shadowing task of two levels of difficulty: slow shadowing (one random number pair per second, or an average of 5.88 bits per second) and fast shadowing (1.5 number pairs per second, or 9.13 bits per second as actually measured). Results indicated that the group performing slow shadowing with tracking showed no decrement in either task. The fast

shadowing group showed significant decrement in both shadowing and tracking. However, the author stated that tracking information lost by the added shadowing was more than offset by information gained in shadowing, a finding that does not fit with a limited channel capacity model.

Cliff did not investigate the mechanism that apparently allowed tracking combined with slow shadowing to go on unimpeded. He merely stated that the two tasks required "a minimal level of joint task difficulty before a performance decrement attributable to their concurrent performance was noted on either task" (1973, p. 247).

It could be questioned here if the result of the slow shadowing experiment could not be interpreted as a sample of parallel processing, but the author did not pursue this matter any further.

The investigative technique for dual-task performance problems is not yet fully developed and is sorely lacking in several areas, according to Welford (1971). One such problem area is the question of the stability of capacity. It has been assumed that the operator's basic capacity is stable and unchangeable, regardless of the addition of a secondary task. Welford (1971, p. 135) cites several studies which tend to dispute this notion, and raises the question of some capacity increase due to the pressure of

the task or to more exerted effort, which would lead to greater "channel capacity."

One of the implications that can be drawn from some of the theoretical models that deal with processing capacity is that the human organism has a flexible system with a variable capacity. As task requirements multiply in demand and complexity, the organism is capable of delivery--up to a point. It is the measurement of this point that has to be dealt with.

Response Modes

In man-machine interface, the operator, once having received the input, must make a decision as to response selection and how to communicate this decision to the equipment. The mode chosen for communicating output to equipment depends upon the particular circumstance. Controls come in various forms: control sticks, key-press, switches, dials, manual or voice control. Experiments on information transduction rate of the human operator have indicated that the man-machine coupling often becomes a bottleneck. This would hinder the operation instead of facilitating and improving the performance of systems control. "Intended human output information must be accepted by equipment which mates the human characteristics"

(Fogel, 1967, p. 392).

Vocal Response Mode

Most studies generally use verbal response in order to test recall, retention, memory, etc., but only a few studies have used vocal response as an alternative for manual manipulation in discrete and concurrent tasks.

One of the earlier experiments in vocal data entry using an experimental acoustic recognition system was reported by Braunstein and Anderson (Note 4) who conducted a single task study involving keypunch operators. Five untrained subjects were instructed to either read the digits aloud or keypunch the digital entries. Even after several hours of practice, the subjects could only keypunch at the rate of 1.0 to 1.5 digits per second (as compared with approximately 2.8 digits per second for a trained operator) while the vocal entry was twice as fast: 2.5 to 3 digits per second. The error rate was the same for keypunching as for vocal data entry. However, the subjects considered keypunching an easier and less tiring task than the vocal entry. The authors concluded that voice input did not seem to offer any speed or accuracy advantages, and that vocal entry of digital data appeared to be preferable only in cases that called for small and infrequent amounts of data

entry. It could be observed here that operator aversion to voice output might be explained by the fact that keypunching allows the operators to converse, eat, or smoke while keypunching, but with vocal data entry such diversions would be impossible.

Alluisi and Muller (1958) compared the information transmission rates of manual and verbal responses to numeric and non-numeric visual codes in a choice reaction time paradigm. Seven different symbolic codes were used with ten symbols in each group of codes. The subjects were required to respond by key-press in the first experiment and verbally in the second experiment.

Verbal response to the numeric codes kept pace with the highest input rate of six bits per second, while the manual response mode had a higher error rate and a lower transmission rate of approximately three bits per second. The vocal superiority appears to be due to its inherent compatibility factor of over-learning of numeric responses, but even with non-numeric codes, the verbal responses showed greater accuracy, albeit slower speed than the motor responses. Generally, "the mode of response apparently had an effect upon performance that was independent of the stimulus; e.g., the best motor-response performance was about equal to the worst verbal-response performance" (p. 253).

One of the few investigations dealing with the vocal response as a viable alternative to traditional response modes was reported by Welch (Note 5). In a comparison of a voice input device with two manual entry devices, Welch used a typewriter keyboard and a pen-and-tablet combination. Other variables of interest also included simple and complex data entry tasks, visual and voice synthesis feedback, practice with the data entry devices, and a two-hand loading task. The results in general indicated that for simple tasks requiring a subject to copy numeric data--three or ten characters--from a CRT rapidly, the keyboard entry method was significantly faster and more accurate than the vocal mode. The keyboard mode was faster for alpha-numeric data entry as well, but was less accurate than vocal entry mode. Vocal entry was superior to both of the manual modes in the performance of complex text data entry, particularly with inexperienced subjects. The vocal mode was faster than both of the manual modes when the subjects were required to perform a two-hand button-holding task during the data entry trials. In many instances, the requirement to correct the voice recognition errors in the vocal mode condition led to additional errors which further reduced the speed of response. Combining CRT feedback with a voice synthesis unit

did not improve speed or accuracy of any type of entry mode. The voice feedback slowed the voice entry mode because some of the subjects waited for the feedback to cease before entering the additional data.

Tracking

Research of motor skills before 1945 concerned itself with outcome measures of performance and showed minimal interest in task variables or response modes. Not until World War II and the appearance of new man-machine systems, did engineering psychologists seriously probe the question of human performance capabilities in complex tasks. The urgent need then for practical solutions called for practical answers to problem-oriented questions.

The task of tracking has often been used as a primary task in investigations of divided attention in dual-task performance and frequently used in conjunction with a verbal response requirement. A distinction must be made here for the sake of clarity of this study: a verbal response is commonly used to illustrate the interference with the tracking task, while the execution of a vocal response signifies here that the voice is being used as a tool or a device as an alternative response mode to investigate the problem of operator overload.

A series of experiments on the effects of digit encoding, retention and/or recalling as secondary tasks on a pursuit tracking task were investigated by Trumbo and associates (Noble, Trumbo, and Fowler, 1967; Trumbo, Noble, and Swink, 1967; Trumbo and Milone, 1971). Tracking was paired with the tasks of either predicting a series of numbers, or with the task of responding to clicks by generating numbers or with the simple task of repeating the numbers. Since the two first mentioned tasks involved choice of response, the result showed more interference than in the task of repeating. In the case of overt verbal response, the result indicated that verbal response does interfere with the primary task of tracking performance, but in the case of learning without overt verbal response, there was no significant interference. In the conclusion of the Noble, Trumbo, and Fowler study it was stated that

> When no overt response selection is required for one of the tasks (Cond. NR), two simultaneous inputs are processed without any apparent decrement in the performance of either task. Put another way, information from two channels can be processed simultaneously and efficiently, provided that one input is effectively stored for overt response at a later time (p. 149).

This conclusion has been disputed by McLeod (1973) who examined the results and found that the subjects, who were supposed to learn a verbal sequence silently without giving overt verbal response, did in fact fail to learn the task, and therefore McLeod considered the result invalidated. He conducted a similar experiment of subjects processing verbal material without overt response in a primary tracking task and found that interference occurred as predicted.

A time-sharing paradigm has been utilized by Johnston, Greenberg, Fisher, and Martin (1970) in a series of experiments that measured the interference effects of verbal memory tasks--encoding, retention, recall--on continuous compensatory tracking performance. In the encoding experiment, word lists were vocally presented to the subjects of 0, 8, or 32 words to be learned during a 16 second trial while concurrently tracking. Recall was tested 4 seconds after the tracking task ended. Results indicated that tracking error increased with higher word presentation rates.

In the retention trial, the subjects were first presented with the word list (0, 1, 2, 4, 6, and 8 words retention load) after which followed a 21 second retention-tracking period. The recall results again indicated that the tracking error increased with the larger

word list.

Preceding each memory recall trial, each subject was given a list of five words. During the tracking segment the subject was required to recall the word list in one of the following ways: free, backwards, alphabetical, or in presentation order. The tracking error was significantly greater for the recall groups than it was for the control groups. Tracking error was the highest in the vocal recall condition and the least in the retaining condition.

All three tasks employed different word lists and were qualitatively different in nature. However, the experiment does provide further evidence that dual-task tracking performance is a function of verbal load.

Damos and Wickens (1977) have recently reported an experiment investigating dual-task performance of a choice reaction time task performed concurrently with a tracking task. The purpose of the experiment was to determine

> . . .which aspects of information processing are affected by the presence of secondary tracking task. This was accomplished by comparing the three parameters associated with the Hick-Hyman Law under single-task and dual-task conditions (p. 211).

The experimental variables were defined as follows: In the choice reaction time task the subjects were required to monitor a cathode ray tube (CRT) display for the presentation of a numeric stimulus to which they responded with the left hand on a key set (3 x 3 digits). There were three levels of stimulus-response alternative sets (the digits: 1-2, 1-4, 1-8) which corresponded to 1, 2, and 3 bits of information. The tracking task required the subjects to nullify the error cursor over the target in a horizontal track on a CRT with corresponding right-left manipulation of a control stick. The cursor was displaced by a random-appearing forcing function. In the dual-task condition, subjects were required to perform both digit entry and tracking as equally as possible. The dual-task display was presented in two different formats. In both conditions the track was located directly above the digit stimulus. In the adjacent condition, both track and digit subtend a vertical visual angle of .8 degrees, while in the separated condition they subtend an angle of 3.8 degrees with an eye-to-display distance of 114 cm. Each subject was tested under both of the single-task conditions and one of the dual-task conditions of adjacent separated displays.

The results are presented graphically in figures in which the average reaction time for the single
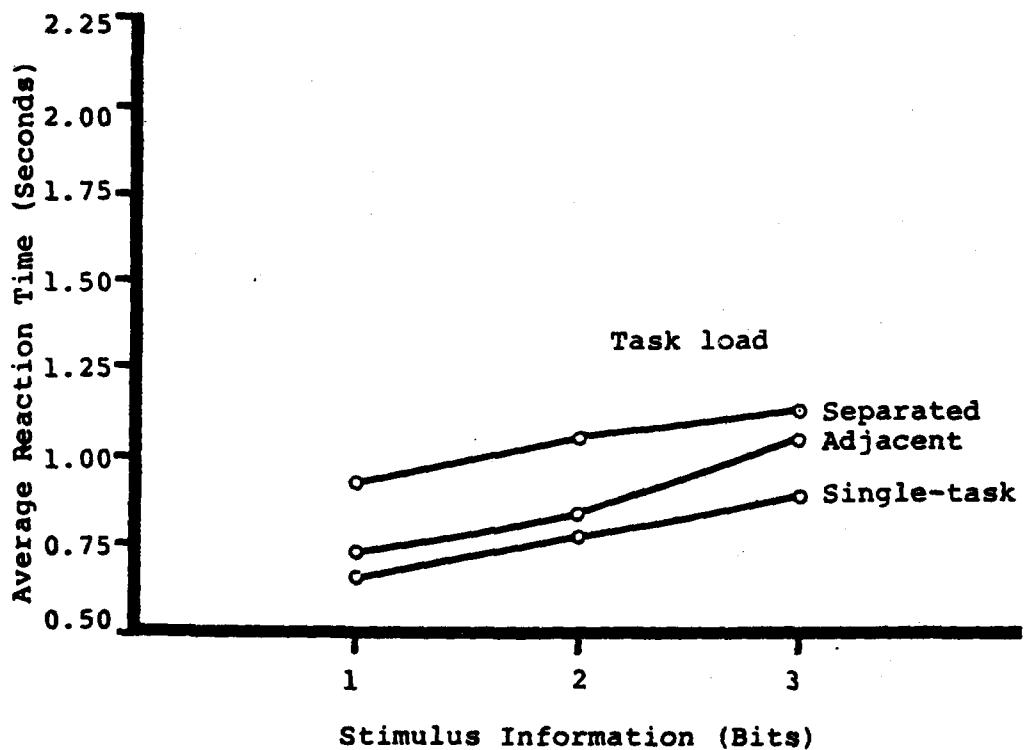
Figure 1.   Results obtained by Damos and Wickens (1977) show-
ing average reaction time as a function of stimu-
lus information and task load in a digit entry
task.

adjacent, and separated conditions are plotted at each level
of stimulus information. Best-fitting linear equations for
the three above groups were also calculated and plotted with
average reaction time as a function of bits of information.
Pairwise comparisons were made for the three equations with
respect to intercept, slope, and linear goodness of fit. It
was reported that the single-task and separated groups as
well as the adjacent and separated groups differed reliably
only with respect to the intercept. The adjace t and sin-
gle-task group equations did not differ reliably with
respect to either slope or the intercept. It was reported
that the linear term for all three data sets reliably
accounted for a significant proportion of the variance while
the quadratic term did not.

The relationship of stimulus-response alternatives,
display separation, and time-sharing on performance of the
tracking task is graphically depicted in Figure 2. A
two-factor mixed model analysis of variance was performed on
the tracking RMS error scores from the last two minutes of
the dual-task trials. The results indicated that while
there were no reliable differences for display separation,
there was a statistically reliable difference in RMS error
score at some level of the stimulus-response alternatives.
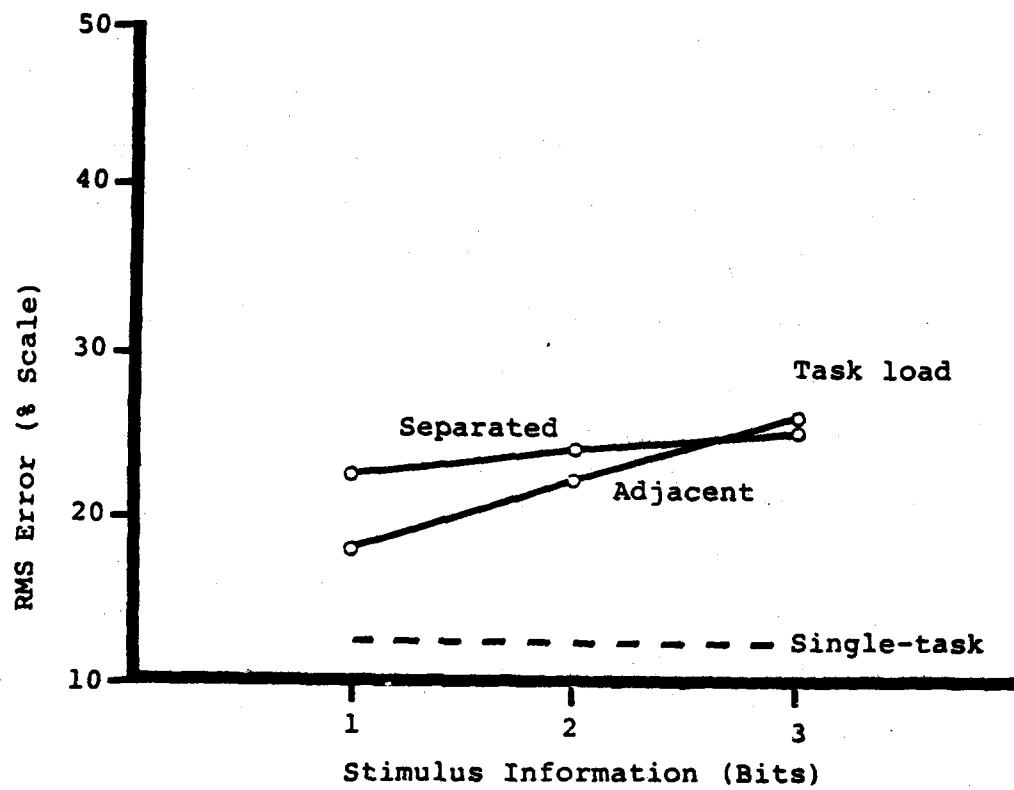
Figure 2. Tracking RMS error as a function of stimulus infor-
mation and task load in tracking task (after Damos
and Wickens, 1977).

Response Alternatives

How does the performance of a task vary when the task to be performed becomes more complex? F.C. Donders, a 19th century Dutch physiologist, (See Fitts and Posner, 1967; Koster, 1969; Welford, 1971) was the first to report that there is a relationship between the uncertainty and the speed of a response. Donders devised three types of experiments to test this hypothesis: Type-a was a simple single stimulus and single response situation in which the time to react was measured; Type-b reaction involved the separate reaction to one of five stimuli with one of five corresponding responses; in the Type-c situation a subject was presented with one of five possible stimuli, but was only required to react to one of them. Donders proposed that each type consisted of separate and similar components which could be subtracted to yield component reaction times. For example, since Type-c has an identification component, and Type-a does not, then Type-a can be subtracted from Type-c, and the result reflects how long it takes to identify the stimulus. Type-b has both identification and selection; the difference between the Type-b and the Type-c times reflects the additional time required to select or choose one out of five alternative responses. Donders' theory did not account for any variation in the separate

components. Each component increased the duration of each type of S-R reaction by a set time, according to Donders.

Hick's Law

Hick (1952) analyzed the outcome of the choice reaction time of Donder's Type-b experiment using information measures. He constructed a model of reaction time performance that is a function of the logarithm to the base 2 of the number of choices in the form of:

$$RT = k \log_2 (n + 1) \qquad (1)$$

where n is the number of stimulus alternatives. When $n = 1$, the k constant is equal to simple reaction time. In testing his formula, Hick also conducted an experiment in which the subjects were presented with a circle-like arrangement of ten lights and a corresponding row of ten telegraph keys on which the ten fingers rested. Less than ten lights were utilized in some conditions. The frequency of occurrence of the lights was balanced, and the order of stimuli was randomized. The resulting data yielded an excellent fit for Hick's model.

Hick's Law has been verified by Hyman (1953) for a similar situation in which he manipulated the number of alternatives, the probability of stimulus occurrence, and

the probability of one stimulus following another. Hyman restructured Hick's formula in a much simpler form:

$$RT = a + b \log_2 n \qquad (2)$$

Stimulus-Response Compatibility

Hick's Law reflects the finding in many choice reaction time studies that as the number of alternatives is doubled, the reaction time is not doubled but increases at a lesser rate which is a function of the logarithm of the number of alternatives and a constant factor--the slope, k in Equation 1, and b in Equation 2. The slope of this function is, according to Hick (1952), the rate of gain of information. The inverse of the slope indicates the information transmission rate dimensioned in bits per second. This index has been used by various writers to compare the effectiveness of various stimulus-response combinations (Fitts and Posner, 1967; Alluisi, 1970; Welford, 1971). It can be seen that as the slope decreases, the information transfer rate increases inversely with the higher degrees of stimulus-response alternatives. This index has been used to compare the efficiency of verbal versus manual response modes for the same input task.

Brainard, Irby, Fitts, and Alluisi (1962) in a

factorial experiment compared information transmission for the following factors: vocal versus manual response modes; visual presentation modes (numerals versus lights); three levels of response alternatives, and self-paced or discrete stimulus presentation. In the first part of the experiment, the S-R pairings were: a) key-press response to correspondingly arranged lights; b) vocal response to numerals on projection screen--these pairings were predicted to show high compatibility. In the second part the pairings were reversed: Key-press response to the numerals; vocal response to the lights--a condition predicted to be less compatible.

The investigators reported that there were no significant differences between self-paced or discrete trials; however, the information loading effect of response alternatives was significant. The results indicated that there was an interaction between stimulus mode and response mode. The information transmission rate for various S-R pairings indicated that some pairings were more compatible than others. The highest rate of 90.9 bits per second was obtained for the combination of vocally naming numerals projected on a screen, while the lowest rate (4.9 bits per second) was reported with naming numerals corresponding to a row of lights. The second highest rate was obtained from

the manual pressing of keys to corresponding lights (9 bits per second). This rate was higher than the rate obtained from pairing the keyboard with projected numbers (5.5 bits per second).

An experiment that illustrates the high compatibility of the perceptual and motor factors that is necessary to yield the highest S-R information transfer rates was reported by Leonard (1959). He found very low error rates and reaction times (less than .2 second) in a series of choice reaction time tasks in which the subjects responded to the vibration of the same key that their fingers rested upon. While the simple RT condition was faster than the choice RT conditions, the slope of this reaction time function was flat or near zero from two alternatives through eight alternatives.

While the above study illustrates the direct relationship of S-R compatibility to the rate of information transmission, other researchers have found that the relationship is specific to each situation. Fitts and Seeger (1953) compared three different stimuli arrangements in a choice reaction time task to three different response arrangements. They reported that there was no one best stimuli code for all response conditions.

Practice

Fitts and Seeger (1953) also evaluated the effects of practice on S-R code compatibility. They found that after three months of 25 training sessions there was an improvement in information transmission efficiency, but the original S-R performance relationships still held.

Practice or familiarity, however, can under certain conditions overcome the S-R mismatch due to increased numbers of response alternatives. Extended practice on several choice RT tasks have yielded results similar to that of Leonard (1959).

Mowbray and Rhoades (1959) reported an experiment in which a subject practiced a two and four choice task for six months. A light panel was used to signal the subject to respond to a similarly arranged keyboard upon which the fingers rested. After some 36,000 trials the RT function slope was zero; the two choice RTs equaled the four choice RTs. Mowbray (1960) also studied the vocal response mode with a different group naming two, four, six, eight, or ten single Arabic numerals in a choice reaction time task. After extended practice, the slope was flat and the intercept--a (Equation 2)--was slightly less than .4 second.

Seibel (1963) trained subjects on a choice RT task with a similar keyboard and light panel as employed by Mowbray

and Rhoades (1959). The ten-finger keyboard, however, was pressed in every possible combination for a total of 1023 alternatives. After some 75,000 responses the average reaction time was approximately .4 seconds, twice the RT of the manual RT of Mowbray and Rhoades (1959) and equal to the vocal RT of Mowbray (1960).

It can be seen that the slope depends on two factors: practice and compatibility. Welford (1971) states "We can therefore identify the steepness of the slope as due largely to the involvement of the translational mechanism" (p. 87).

Statement of the Problem

The appearance of voice recognition devices has given an added dimension to the study of human performance in time-sharing environments and opened up new possibilities for a reduction in the amount of controls that an operator has to cope with. In a man-machine interface where the design could allow for a voice response rather than the traditionally used manual controls, the question of the effectiveness of a vocal response mode must be investigated.

Previously cited authors (Hill, 1972; Feuge, Note 2) have hypothesized that vocal response would be independent of other channels; would interfere less with other activities; would be faster than manual response; and would

be more "natural" for the human operator than other modes.
These assumptions have not been formally tested and the
question of how vocal data entry compares with manual modes
in varying degrees of task work load has not been answered.

The "vocal channel" can be said to be independent of
the "manual channel" if activity in one does not disrupt
activity in the other. This parallel processing view has
been rejected by several researchers reviewed in the
preceding section. The common element in most of this
review was that processing dual tasks of sufficient
difficulty is likely to lead to decreased performance for
one or both of the tasks with respect to both of their
isolated performances. This single-channel concept
(Broadbent, 1958; Welford, 1971, and others) holds that the
performance interference in processing information from
multiple sources is due to the need for the information to
share a common mechanism at certain processing stages. This
mechanism is of limited capacity and can only process infor-
mation in series. When two tasks require processing from
the same mechanism, interference in the performance of one
or both tasks occurs. Most of the time-sharing paradigms
reviewed in this paper report this interference effect for
various combinations of tasks, stimuli, and response
modalities.

In several experiments involving word memory tasks with vocalization and tracking (Trumbo, et al, 1967; Johnston, et al, 1970, and others), the manual tracking performance showed an interference effect. However, the effect on verbal performance is impossible to derive from these experiments. In a dual-task study reported by Damos and Wickens (1977) involving a digit entry task, both manual digit entry and manual tracking performance demonstrated a decrement. The question of whether these findings will hold for a vocal entry task, as well, is of primary concern to this study.

The independence of vocal and manual concurrent entry performance could not be supported in the earlier reported experiment by Schvaneveldt (1969).

The single-channel hypothesis would also predict that as the difficulty of one of the tasks in a dual-task paradigm increases, performance on the first as well as on the second task will deteriorate. In the Damos and Wickens study, increasing the information load by increasing the numbers of digits to which the subject had to respond, resulted in a performance decrement for the manual digit entry mode as well as the manual tracking performance.

Another goal of the present study is to investigate the vocal response mode under similar conditions as those

employed by Damos and Wickens. Would increased task load, as induced by increasing response alternatives, be directly related to an increase in vocal response latency as predicted by the Hick (1952) or Hyman (1953) models?

The parallel processing hypothesis (Kahneman, 1973; Neisser, 1967, and others) on the other hand would predict that there would not be any decrement in either mode of digit entry performance or tracking performance from the single to the dual task situation. In addition, these performances would not be affected by the load induced by increasing numbers of response alternatives. However, the only convincing evidence of this phenomenon as seen in this review, has been reported for experiments where the two tasks were very easy (Mowbray, 1953; Cliff 1971, 1973); where the response dealt with overlearned items (Alluisi and Muller, 1958); where the subjects had had extended practice (Mowbray and Rhoades, 1959; Mowbray, 1960); or where stimulus and response were highly compatible (Leonard, 1959).

The purpose of the study was to investigate vocal and manual digit data entry under single and dual-task conditions. Tracking was performed singly as well as in combination with either vocal or manual response modes. The response alternative sets of four, eight, or sixteen digits (from 1-16) were factorially combined with the above conditions.

A performance feedback method (Gopher and North, Note 7; Norman and Bobrow, 1975) was utilized to allow the subjects to attend to both tasks with equal priorities. This was designed to prevent the subjects from concentrating on the harder task, as in the Mowbray (1952, 1953) experiments.

The above experimental paradigm would allow for the previously mentioned comparisons related to vocal, manual and tracking performance decrements with increasing load to be made. The factorial design of the experiment would also allow contrasts to be made between the vocal and manual response modes in order to determine which mode interferes less with tracking and which is superior in speed and accuracy.

Hill's (1972) suggestion that the vocal response mode interferes less with the performance of other activities was tested in this study. The tracking performance associated with the vocal and the manual modes was compared in order to determine which mode has the higher interference effect on the concurrent tracking task. Similarly, the effect of the three levels of response alternatives associated with the two response modes on dual-task tracking performance was compared.

Previous findings of speed and accuracy of the vocal response mode have not been conclusive. While some single-task studies have shown that the vocal response mode was as fast as the manual mode and had a lower error rate (Braunstein and Anderson, Note 4; Alluisi and Muller, 1958; Brainard, et al, 1962), voice entry using an isolated word recognition system was at a disadvantage compared with keyboard entry in simple tasks, according to Welch (Note 5). The Welch study showed voice entry to be superior in complex tasks, while the Schvaneveldt (1969) study found vocal task performance to be lagging in time as compared with the manual performance.

The unresolved problem of vocal or manual response mode superiority in time-sharing, emphasizes the need for controlled investigations. By obtaining performance estimates for both vocal and manual modes under all treatment combinations the question of which modality is faster or more accurate could be addressed. The above review suggests that it is highly unlikely that one mode of response is superior to the other for all treatment combinations. The interaction effects of the proposed variables is not known, but it is of great interest to this study.

The hypothesis that the vocal response mode is more natural for information transmission was related to the

findings in the review of the effects of compatibility and practice on performance.

Fitts and Seeger (1953) found that stimulus-response compatibility is specific to each situation. Brainard, et al, (1962) reported that the vocal mode was faster for digit entry tasks, while the manual mode was faster in digitizing a row of lights. The vocal interface as tested in this study takes advantage of a unique situation somewhat similar in nature to that of Leonard (1959), who found extremely low reaction times and error rates for a highly compatible vibrating keyboard. Several other studies have shown similar effects for both the vocal and the manual response modalities after extended practice (Mowbray and Rhoades, 1959; Mowbray, 1960; Seibel, 1963).

Even though it was beyond the scope of this study to test the "naturalness" of either response mode, both S-R compatibility and practice are factors which seem to favor the vocal response mode.

The highly familiar digit entry task combined with the vocal response mode was hypothesized to contribute less of a performance decrement from the single to the dual task situation for the vocal response mode. Familiarity with the vocal task also suggested that interference with tracking could be attenuated and that increased response alternatives

could have a minimal effect on the vocal and tracking performance. The same factor could make the vocal response mode faster than the manual as well.

The S-R compatibility factor should depress the manual response mode performance in several aspects. In the manual mode there will probably be a tendency to attend to the keyboard which will interfere with both the response rate and the tracking performance.

Task Background and Selection Rational

The two tasks that have been selected for inclusion in this experiment are digit entry and compensatory tracking. These tasks have been employed in both single and dual-task research. The following is a description of each task, its past employment, and the reasons for inclusion of the tasks in the present study.

Digit Entry

The digit entry--or choice reaction time--task has been in continuous use in experimental psychology for well over a century. Research associated with the choice reaction time task has centered around the variable of the number of choice alternatives. Hick (1952) and Hyman (1953) have developed predictive models that describe reaction time performance as a function of the number of alternatives. The

task-loading effect of response alternatives has been
verified in many situations. The effects of compatibility
have been reported regarding choice reaction time experi-
ments (Fitts and Seeger, 1953) as well as practice (Leonard,
1959).

The digit entry task employs the visual modality for
the presentation of the digits on a CRT display. Recent
investigators employing this presentation mode include Damos
and Wickens (1977), and Welch (Note 5). The digit display
is presented in this study with a tracking task during
time-sharing task performance. Both the displays are
presented in such a fashion that they are in full view with
a single foveal fixation point. Damos and Wickens (1977)
have demonstrated that a display with a separated track and
digit presentation has a significant interfering effect on
data entry performance. The visual rather than the auditory
presentation mode has been selected in order to avoid the
"slowing" effect of auditory presentation (Welch, Note 5).

The vocal response mode has been utilized by several
researchers reviewed in this report, but only Welch (Note 5)
has utilized a voice recognition device rather than a voice
key. The voice device has an advantage in that it is
directly interfaced with the experimental apparatus and is
more reliable than an experimenter performing timing and
scoring operations. However, since the device requires a

precise vocalization, a poor vocalization was recorded as an error, just as a poor keypress would be. These machine errors could induce an additional load on the subjects if they were required to correct them (Welch, Note 5). Therefore, errors were not correctable.

The manual response mode has been the most common mode utilized in choice reaction time experiments. Experimenters have employed various devices to time the manual response. Seibel (1963) and Leonard (1959), for example, have employed special ten-finger keyboards for extremely high speed data entry. Other have used standard correspondence typewriter keyboards (Welch, Note 5) or IBM card punch keyboards (Braunstein and Anderson, Note 4).

The manual data entry device selected for this study was a sixteen key keyboard arranged in a four-by-four matrix. It was assembled with military-industrial quality switches, incorporating a mechanism which prevents actuation of adjacent switches. The numerals on the switches are backlit. The force required to actuate these switches is higher than the force typically employed in the laboratory. Could these switches introduce a loading effect due to the high key force? Galanter and Owens (Note 6) have investigated this problem. They performed a simple reaction time study in which the influence of three different levels of

key activation force was evaluated with respect to the distributions of reaction time responses. It was reported that the effect of key force was not significant.

Tracking

Tracking is a highly demanding task that has been used only within the last three decades to evaluate the coordinated manual-visual motor skill behavior in the study of human performance. Tracking has been of interest in time-sharing performance as well. It has been found to be sensitive to the effects of interference due to concurrent performance with several tasks. Some investigators have used tracking to evaluate memory processes in which there is no direct memory performance measure available (Noble et al., 1967; Trumbo and Milone, 1971; McLeod, 1973). Continuous compensatory tracking has been used to measure the interference effects on information processing due to a secondary tracking task (Damos and Wickens, 1977).

The tracking task selected for this study was developed by Gopher and North (Note 7). It is a continuous compensatory tracking task in which a subject is required to adjust a control stick in order to return a cursor to the center of a horizontal track.

Time-sharing Tasks

Many writers have pointed out the need to allocate attention among several tasks for effective time-sharing performance (Kahneman, 1973; Norman and Bobrow, 1975; Navon and Gopher, Note 3). Gopher and North (Note 7) have developed a technique that presents the subject with an effective feedback display on which he can monitor his performance on two tasks being time-shared. Adjustments can then be made, if possible, so that the subject is performing each task with equal priorities.

CHAPTER II
METHOD

Subjects

The subjects were seventeen volunteers from a subject pool available to the Naval Aerospace Medical Research Laboratory in Pensacola, Florida. The subject population was comprised of two male Naval aviation students, three female college students, and twelve male college students. The college students were paid for participating in the experiment.

The subjects' ages ranged between nineteen and twenty-six years. Every subject was right-handed, and was screened for visual defects and speech impediments. Every subject completed the three hour experiment, and every subject was included in the data analysis.

Apparatus

The experimental control system utilized in this study was the Multipurpose Automated Research Test Station (MARTS) located in the Aerospace Psychology Department of the Naval Aerospace Medical Research Laboratory in Pensacola, Florida. The MARTS system has been described by Harris, North, and Owens (Note 8). Figure 3 illustrates the functional

Hewlett-Packard 1310A

CRT Display

Megatek 6000 Graphic Controller

Measurement Systems Handcontrol

A/D Converter

Keyboard

Digital I/O Interface

Microphone-Headset

Scope Voice Recognition

Wangco Disk

Data General Nova 800 Computer

Data Media Terminal

Versatec Printer
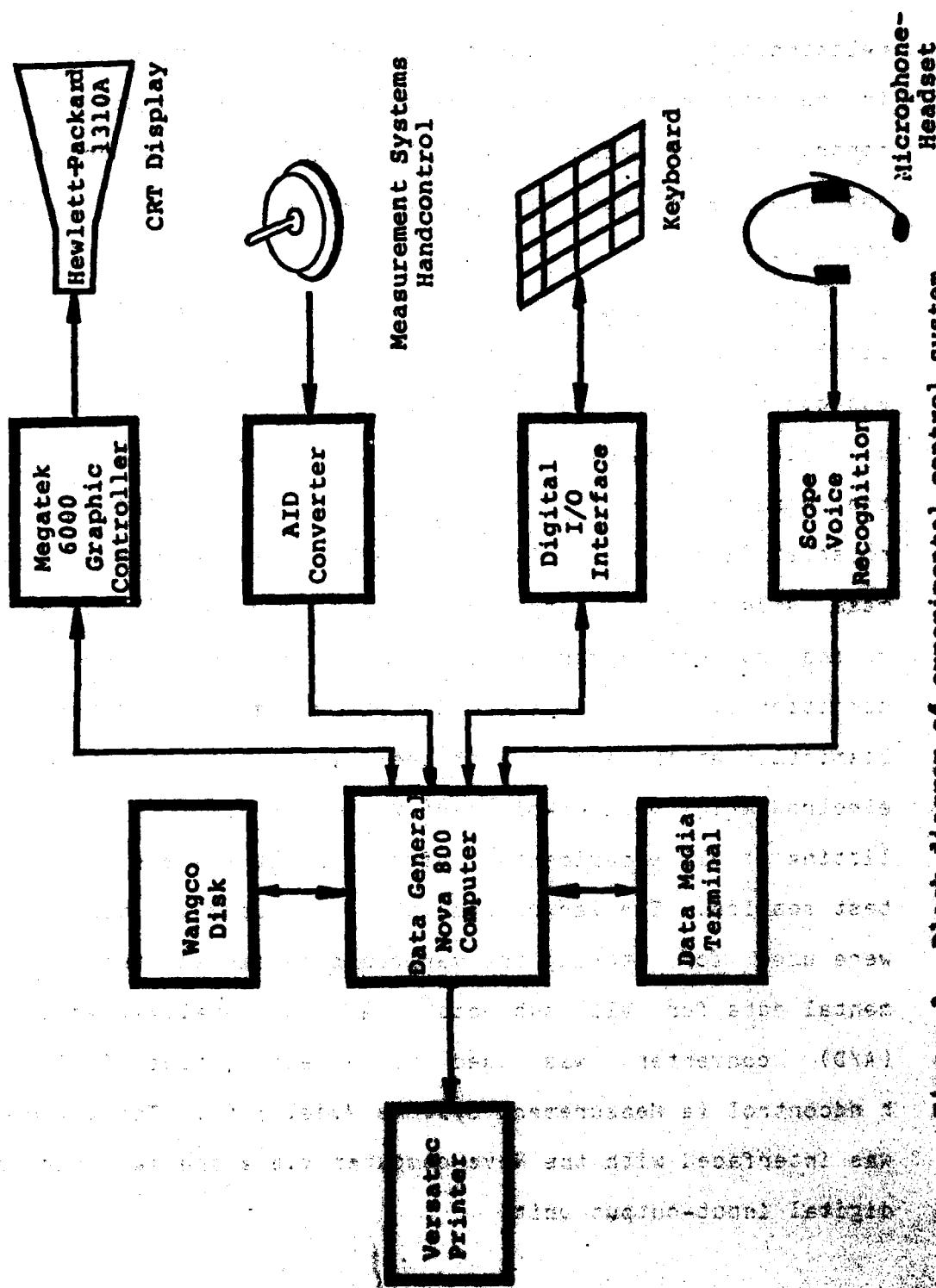
Figure 3. Block diagram of experimental control system.

relationship among the various experimental devices employed in the experiment. The apparatus included equipment for the control and presentation of stimuli, and for the recording and on-line analysis of responses as well as a subject test station.

The following description of the MARTS system is outlined in Figure 3. The experimental devices were interfaced and controlled by a Data General Nova 800 computer with 32K x 16 bit core memory. Peripheral devices used directly with the Nova minicomputer included 1) a Data Media CRT terminal console, 2) a Versatec line printer, and 3) a Wangco disk mass storage unit. The experimenter used the CRT for entry of the subject parameters, control of the experimental conditions, and for the display of subject performance statistics at the end of each test segment. The Versatec electrostatic printer-plotter was used for the complete listing of all experimental information at the end of each test session. The Wangco removable disk pack and disk drive were used for storing the various programs and the experimental data for all subjects. A Nova analog-to-digital (A/D) converter was used to accept output from the handcontrol (a Measurement Systems Model 526). The keyboard was interfaced with the Nova computer via a special purpose digital input-output unit.

A Scope Electronics Voice Data Entry Terminal System (VDETS) was the speech recognition unit. The Megagraphics 6000 Megatek display generator with a Hewlett-Packard 1310A CRT display presented the dynamic tracking display, numerals and performance feedback. The keyboard was configured with sixteen momentary contact push-button switches. The keys were arranged on a panel in four rows of four keys in each row.

Experimental Design

Two experimental designs, both based on the analysis of variance, were employed in this study. A randomized block factorial design, designated RBF-pq by Kirk (1968), was used in evaluating the effects of three levels of response alternatives (four, eight, or sixteen stimuli in a set) and three levels of tracking task load (tracking alone, tracking with manual entry, or tracking with vocal entry) upon tracking performance (see Figure 4). A randomized block factorial design--RBF-pqr--was also used in evaluating the effects of three levels of response alternatives, two levels of response mode, and two levels of task load upon the performance of a digit entry task. The response alternatives were four, eight, or sixteen digit sets selected from the digits
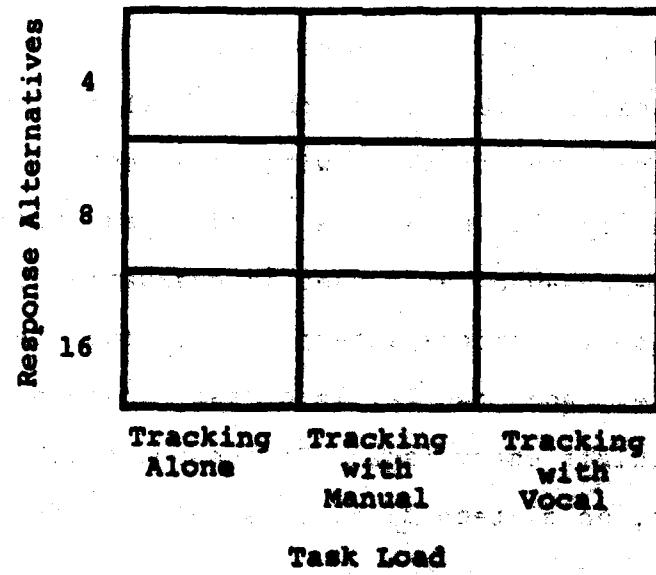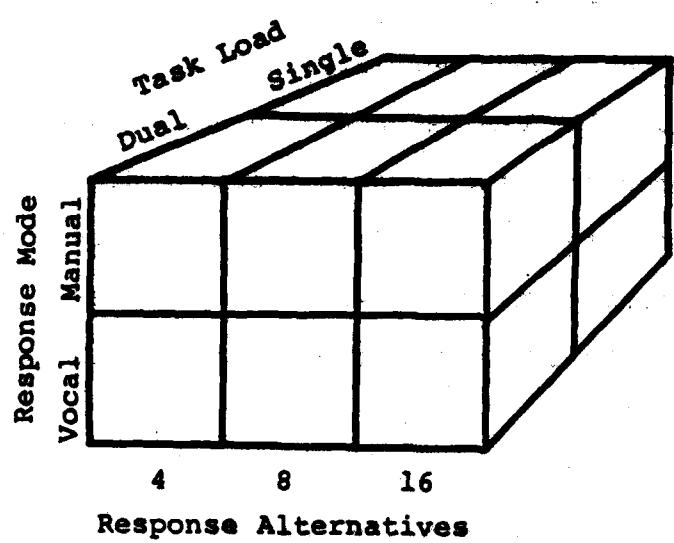
Figure 4. Experimental Treatments.

one through sixteen; response mode was the manual or vocal entry of the digits; task load was the performance of the task by itself or concurrent with the tracking task. The subjects were assigned a random order of block treatments.

Dependent Variables

The dependent variables recorded in this study were: 1) digit entry response latency interval measured from start of digit presentation to the registration of a response to that digit; 2) digit entry errors, and 3) tracking root mean square error.

Procedure

The subjects were initially shown the experimental control system briefly before being seated at the subject test station. An introductory statement was read to the subject by the experimenter (see Appendix for instructions to the subjects). The order of the experimental procedures started with a calibration session, which was followed by a block of five practice trials. The three experimental blocks that followed were also composed of five trials. Subjects were randomly assigned one of six orders of administration of the experimental blocks.

The practice block of trials were started with a sin-

gle-task manual entry of the full digit set. This was followed by a rest and instruction period. The single-task vocal entry task was presented next in a similar manner, followed by a rest and instruction period. Single-task tracking was next, followed by the rest period and the reading of the instructions for the dual-task practice trials. The manual task was combined with the tracking task, which was followed by another rest and instruction period. The practice block ended with the dual-task vocal entry and tracking trial. Another rest period followed, in which the subjects were free to move around.

The experimental trial procedures were similar to the practice sessions, except that within each of the three blocks, it was randomly determined in which order the manual and the vocal single and dual task trials would be administered. The single-tasks always preceded the dual-tasks, and the single-task tracking trial was always the third trial in any block.

Calibration Trial

The subject was fitted with a microphone headset combination. The instructions regarding the keyboard and voice recognition device were read to the subject and the operation of these devices demonstrated. In order for the

voice recognition device to function reliably, it has to be calibrated to each individual subject's voice. The subjects were instructed to repeat the numerals one through sixteen ten times each on cue from the machine. The subjects were given the same amount of practice on the manual keyboard. The voice recognition device was then tested to see if it was "tuned" properly to recognize correctly the test numerals spoken by the subject, while also matching at least 122 characteristics of 128 for the correct word. Failure to reach this criterion necessitated "re-training" of the machine on the poorly recognized words.

Single-task Digit Entry

The digit entry task required the subjects to vocalize a digit or press a key corresponding to a digit presented on a CRT display. The CRT display was mounted directly in front of the subject at approximately eye level at a distance of 28 inches. The digit stimuli were presented from one of several response alternative sets: four digits from the numerals 1 to 4, 5 to 8, 9 to 12, or 13 to 16; eight digits from 1 to 8, or 9 to 16; and sixteen numerals from 1 to 16. There were 128 digits in each list of stimuli. Each stimuli list was randomly generated via a computer program with the restriction that each numeral

appear 32, 16, or 8 times for the four, eight, and sixteen response alternative conditions respectively. A stimulus list was regenerated if the sequence of a single digit repeated itself more than three times. The single-task condition was self-paced. As soon as the subject responded to the presented digit stimulus on the CRT with any response, a new stimulus was displayed. In order to indicate to the subject that a new digit was being shown, a tenth of a second blanking of the digit was inserted between the stimuli.

The subjects were also presented with a feedback display for the digit task above the track and to the left of the digit display as shown in Figure 5. Subjects were instructed to maintain both a high degree of accuracy and speed on the task in order to keep the displayed feedback bar graph as high as possible. The height of the feedback bar corresponded to the mean latency of the subject's last five correct responses. The height of the bar was recomputed and updated after every response. The rectangular goal box on the feedback indicator corresponded to a response latency of 1.5 seconds.

During the manual response mode the subjects were instructed to respond only to the digit stimuli on the screen, and cautioned that any attempt to correct errors
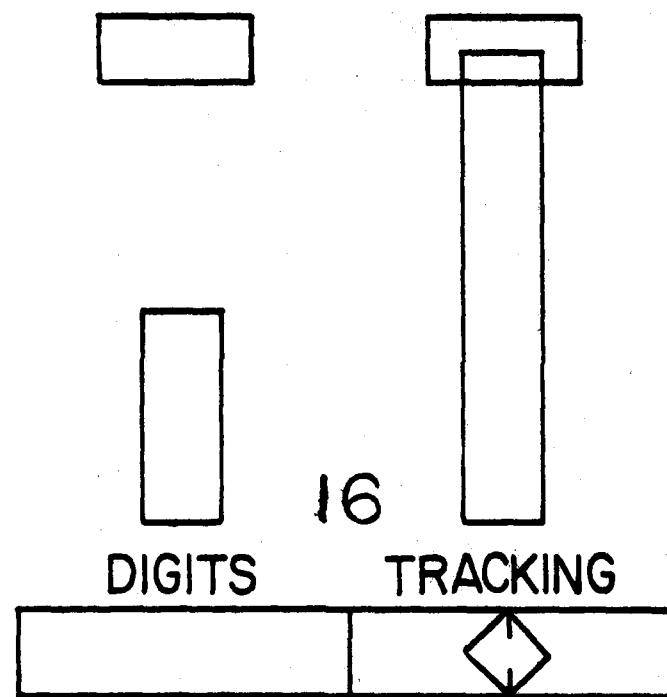
Figure 5. Arrangement of display.

would only result in additional errors. The keyboard was labeled with the digits 1 through 4 from left to right on the top row, with succeeding rows of 5-8, 9-12, and 13-16. The appropriate rows were backlit to correspond with response alternative treatment condition. The subjects were required to use only their left hand for pressing the keyboard. In addition, each subject was instructed to return his fingers to a designated "home" row in the eight and sixteen response alternative conditions.

The vocal response mode procedure was similar to the manual with the exception that the subject was only informed verbally as to which digit set would be presented, since the keyboard would not be backlit. The subject was cautioned not to stutter or attempt to correct his responses, as this would be scored as more errors. Furthermore, if a word was not spoken loudly enough, it would have to be repeated util a response was registered and new stimulus was presented.

At the end of the single-task digit entry trial of approximately 2.5 to 3 minutes, the entire display was turned off and the subject rested approximately three minutes while the next task was being prepared. The experimenter recorded the mean and standard deviation of the correct response latency. The vocal errors were also transferred from the digit stimulus list at this time.

Single-task Tracking

All subjects were required to perform a one-dimensional, compensatory tracking task requiring appropriate left-right movements of a handcontrol to maintain the position of a diamond-shaped cursor on a vertical line in the center of a horizontal track (see Figure 5). A forcing function input in the computer program forced the cursor off the center line. The forcing function consisted of the addition of three non-harmonically related sinusoidal waveforms. To the subjects, the movements of the diamond appeared to be random. The handcontrol employed was a Measurement Systems Model 526 finger control with a lateral deflection range of +/- 30 degrees. The handcontrol was extended to decrease fatigue during performance and allow the subject to control the stick more easily. It had no return spring. The handcontrol was optimally positioned to be operated with the right hand.

A performance feedback indicator identical to the digit entry feedback was presented to the subjects on the right side of the display. The height of the associated goal box represented an average tracking error of 20 percent of the length of the track. The position of the indicator was updated in one second intervals. The computations were based on the last ten seconds of tracking root mean square

error. The subjects were instructed to try to keep the indicator bar as high as possible by fast and accurate responses to the tracking task.

The tracking trial was always the third trial in any block of trials. The duration of the trial was set at two minutes. The CRT display was turned off at the end of the trial and the subject rested for approximately three minutes. During this period, the experimenter recorded the mean and standard deviation of the tracking root mean error for the entire two minute trial. The next trial was prepared for by entering the statistics for the tracking trial and one of the digit entry modes as well as other parameters.

Dual-task Procedures

The fourth and fifth trials in any response alternative block required the subjects to perform the two minute tracking task concurrently with either the manual or the vocal digit entry task. The selection of the manual or vocal task was randomized in order to minimize order effects. The subject was informed which task he was to perform. Preceding the manual task, the appropriate keys were backlit, and the subject was reminded to keep his fingers on his home row. Due to the weight of the headset, the

subjects could not be required to wear it for the entire three hour experiment. Intermittently, it had to be re-mounted and adjusted.

The feedback display included both the digit and the tracking performance indicators as shown in Figure 5. The goal boxes represented the correct digit entry mean latency and the mean tracking error score attained by the subject in the preceding single-task trials. The range of the feedback scale corresponded to 1.5 standard deviations above and below its respective mean. The height of each of the feedback indicators was a standard score function. This was computed by subtracting the current task mean from the single-task mean, and dividing the remainder by the standard deviation of the single-task mean. The digit feedback was updated after every response and its current mean was computed with the last five preceding correct response latencies. The position of the tracking feedback bar was updated every second. The computations were based on the scores of the preceding ten seconds of performance.

The subjects were instructed that the tasks were of equal importance, and that they should attempt to reach or exceed their previous performance level for both of the tasks. They should attempt to keep the two bars even and not sacrifice performance of one at the expense of the other.

At the end of the dual-task trials, the CRT display was again turned off and the subject rested for approximately three minutes. The experimenter recorded the means and standard deviations from each task and prepared for the next trial. If the trial was the last in a block, the subject was offered a soft-drink and the opportunity to get up and walk around during a ten minute rest period. The subject was again seated and informed about the next block of trials corresponding to a response alternative level.

At the end of the experiment, the subjects were thanked and paid for their participation. In a short debriefing session an explanation of the purpose of the study was given. Individual results were discussed and questions were answered.

CHAPTER III

RESULTS

Digit Entry Task

Descriptive statistics for digit entry task correct response latency (CRL) may be found in Table 1. An examination of the table will reveal that there are several noteworthy differences of interest to this study. The analysis of variance procedure as outlined by Kirk (1968) for a randomized block factorial (designated RBF-pqr) design was selected. This analysis is presented in Table 2.

An examination of the correct response latency scores revealed that the scores for each treatment combination were approximately symmetrical in distribution with a slight positive skewness. A logarithmic transformation, $X' = \ln (X + 1)$ was selected. This transform function has been used for reaction time data when the data were positively skewed and the treatment means and standard deviations proportional.

The main effects were all found to be significant. The main effect of response mode was found to be significant with F (1, 176) = 31.623, $p < .01$. The means for the manual and vocal response modes were 1.288 and 1.111 seconds respectively. Figure 6 illustrates the mean correct re-

TABLE 1

DESCRIPTIVE STATISTICS FOR CORRECT RESPONSE

LATENCY IN DIGIT ENTRY TASK

Condition	Mean	Standard Deviation	Variance
A - Response Mode			
Manual	1.288	.4198	.1763
Vocal	1.111	.0386	.0015
B - Task Load			
Single	1.059	.1322	.0175
Dual	1.340	.3625	.1314
C - Response Alternatives			
Four	1.023	.1352	.0183
Eight	1.204	.2243	.0504
Sixteen	1.372	.4211	.1774

TABLE 2

ANALYSIS OF VARIANCE SOURCE TABLE
FOR CORRECT RESPONSE LATENCY

Source	SS	df	MS	F
Blocks	.568	16	.036	5.583**
Treatments	2.685			
A – Response Mode	.201	1	.201	31.623**
B – Task Load	.693	1	.693	108.965**
C – Response Alt.	.731	2	.366	57.517**
AB	.409	1	.409	64.425**
AC	.559	2	.279	43.949**
BC	.057	2	.028	4.457*
ABC	.035	2	.018	2.786 NS
Residual	1.119	176	.006	
Total	4.372	203		

* $p < .05$

** $p < .01$
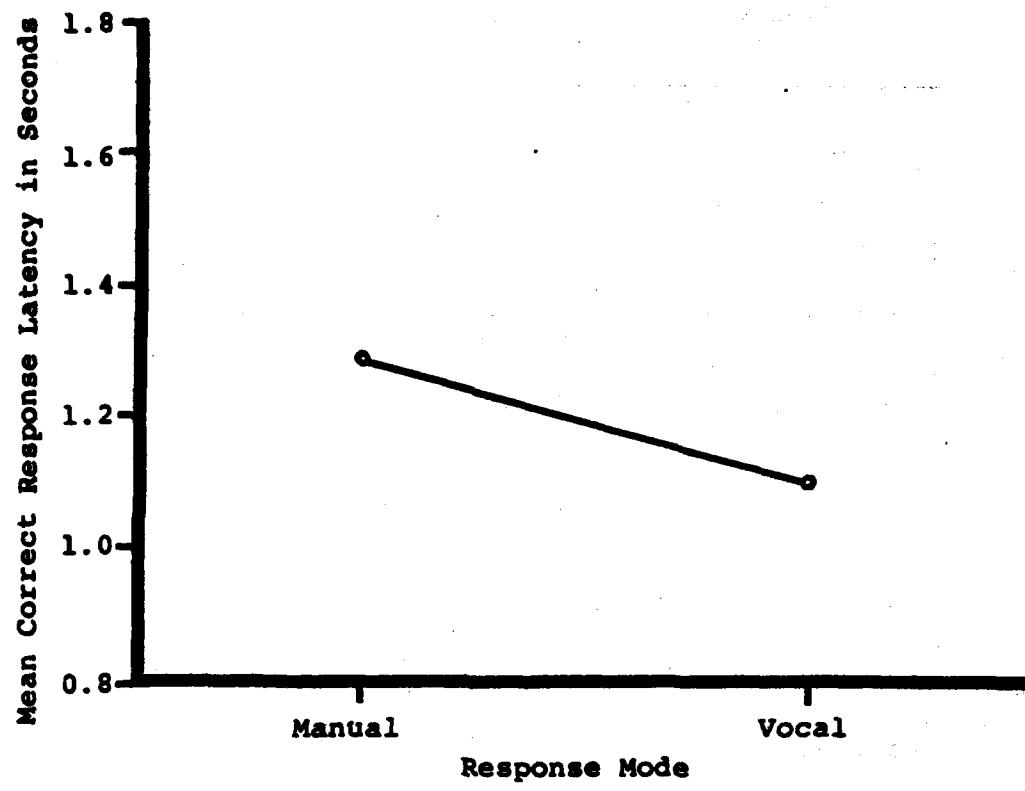
Figure 6.   Mean correct response latency as a function of
response mode in digit entry task.

sponse latency as a function of response mode in the digit entry task. The task load condition was found to have a significant effect on response latency, $F(1, 176) = 108.965$, $p < .01$, with means of 1.059 and 1.340 seconds for the single-task and the dual-task treatment levels. The mean correct response latency as a function of task load is depicted in Figure 7. The variable of information load or response alternatives with mean latencies of 1.023, 1.204, and 1.372 seconds for the four, eight, and sixteen response alternatives respectively yielded a significant main effect, $F(2, 176) = 57.517$, $p < .01$. This effect can be seen in Figure 8, in which the mean correct response latency is plotted as a function of logarithm to the base two (bits) of the number of response alternatives. Since higher order effects were found, the three main-effects are not of primary interest even though they do generally reflect what effect the three independent variables had on the dependent variable.

The interaction of response mode by task load (AB term) was found to be significant with $F(1, 176) = 64.425$, $p < .01$. The response mode by response alternative interaction (AC) was significant with $F(2, 176) = 43.949$, $p < .01$. The task load by response alternatives was also found to be significant with $F(2, 176) = 4.457$, $p < .05$.
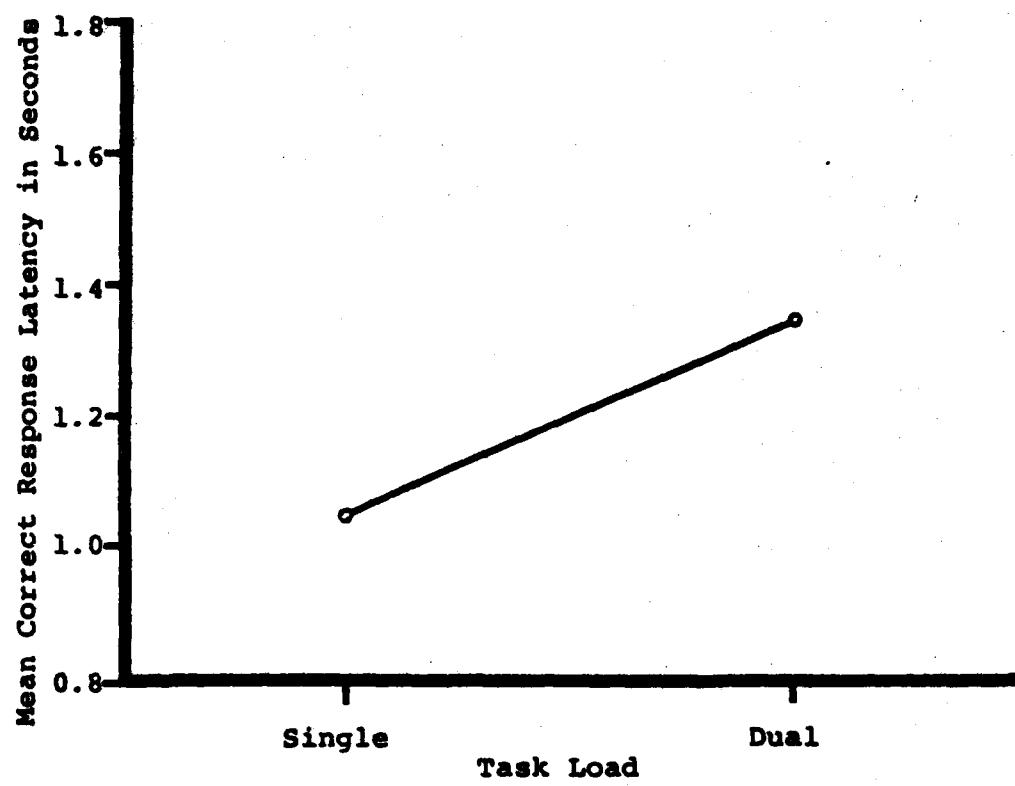
Figure 7. Mean correct response latency as a function of
task load in digit entry task.
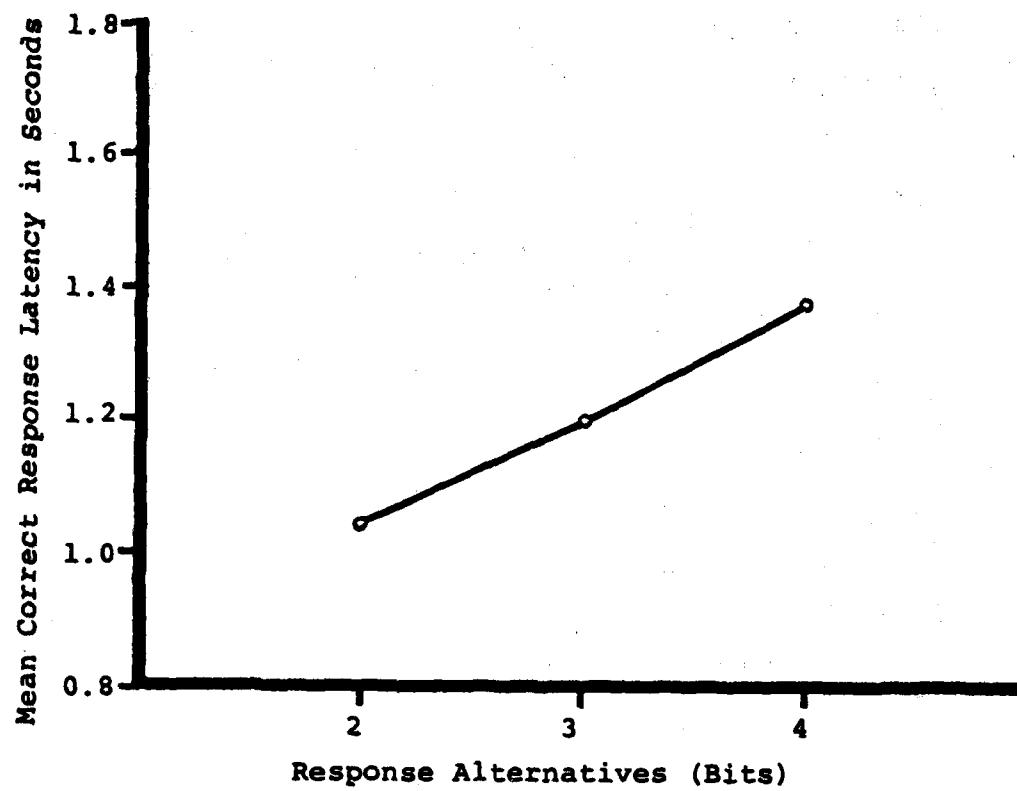
Figure 8. Mean correct response latency as a function of
response alternatives in digit entry task.

However, the second order interaction of response mode by task load by response alternatives (ABC) was not found to be significant. This indicated that an examination of the simple main effects would be of value, while computing statistics for the simple simple main effects and the simple interaction effects would add no additional information.

The simple main effects for the above significant interactions are tabulated in Table 3. All the statistics were found to be significant with p less than .01 with the exception of C at a2, which was not found to be significant. Response mode is shown at the two levels of task load in Figure 9. The manual mode was found to be superior to the vocal mode at the single-task level with means of 1.035 and 1.083 seconds, while in the dual-task condition the vocal mode was faster than the manual mode with means of 1.140 and 1.540 seconds. The manual response mode again was found to be faster in latency than the manual condition at the four response alternative level (.954 vs 1.091), while the vocal mode was superior at the eight and sixteen level (vocal: 1.113, 1.130 vs manual: 1.296, 1.613 seconds). This relationship may be more clearly seen in Figure 10, in which correct response latency is a function of response mode and response alternatives.

The simple main effects analysis indicated that the

TABLE 3

ANALYSIS OF VARIANCE TABLE FOR SIMPLE MAIN EFFECTS

IN DIGIT ENTRY TASK

| Source	SS	df	MS	F
A - Response Mode				
A at b1	.054	1	.054	8.556**
A at b2	1.768	1	1.768	277.990**
A at c1	.174	1	.174	27.338**
A at c2	.194	1	.194	30.483**
A at c3	1.151	1	1.151	181.049**
B - Task Load				
B at a1	3.236	1	3.236	508.950**
B at a2	.069	1	.069	10.919**
B at c1	.184	1	.184	28.889**
B at c2	.494	1	.494	77.649**
B at c3	.817	1	.817	128.477**
C - Response Alternatives				
C at a1	2.561	2	1.281	201.398**
C at a2	.011	2	.006	.867 NS
C at b1	.377	2	.189	29.654**
C at b2	1.191	2	.595	93.627**
Residual	1.119	176	.006	

** $p < .01$
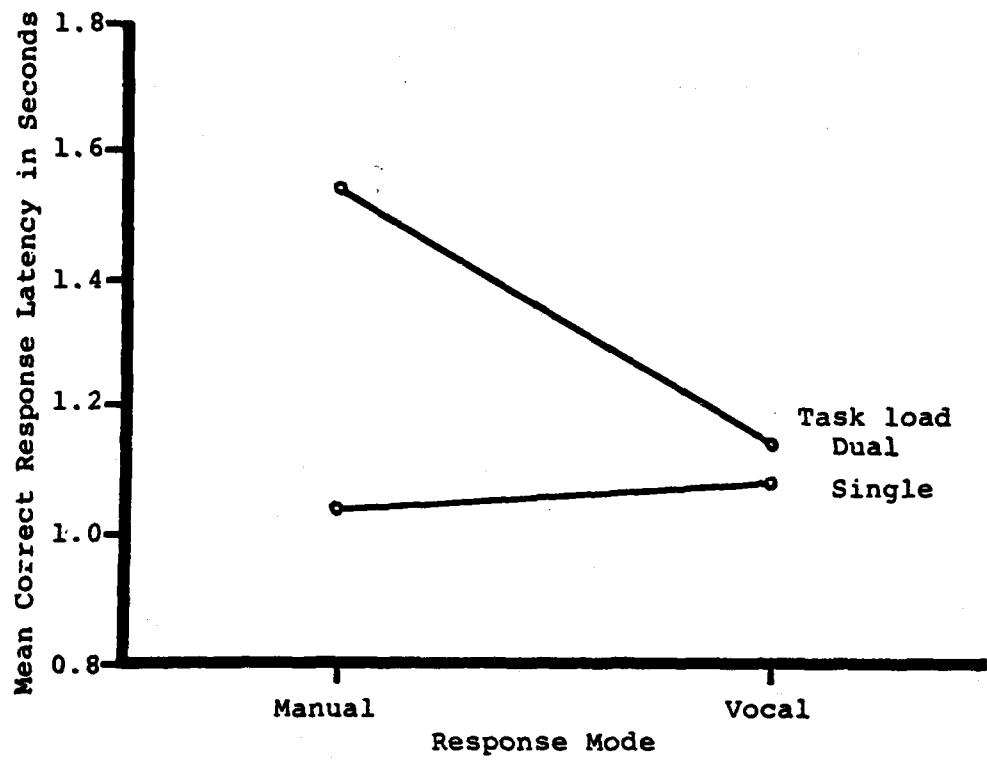
Figure 9. Mean correct response latency as a function of
response mode and task load in digit entry task.
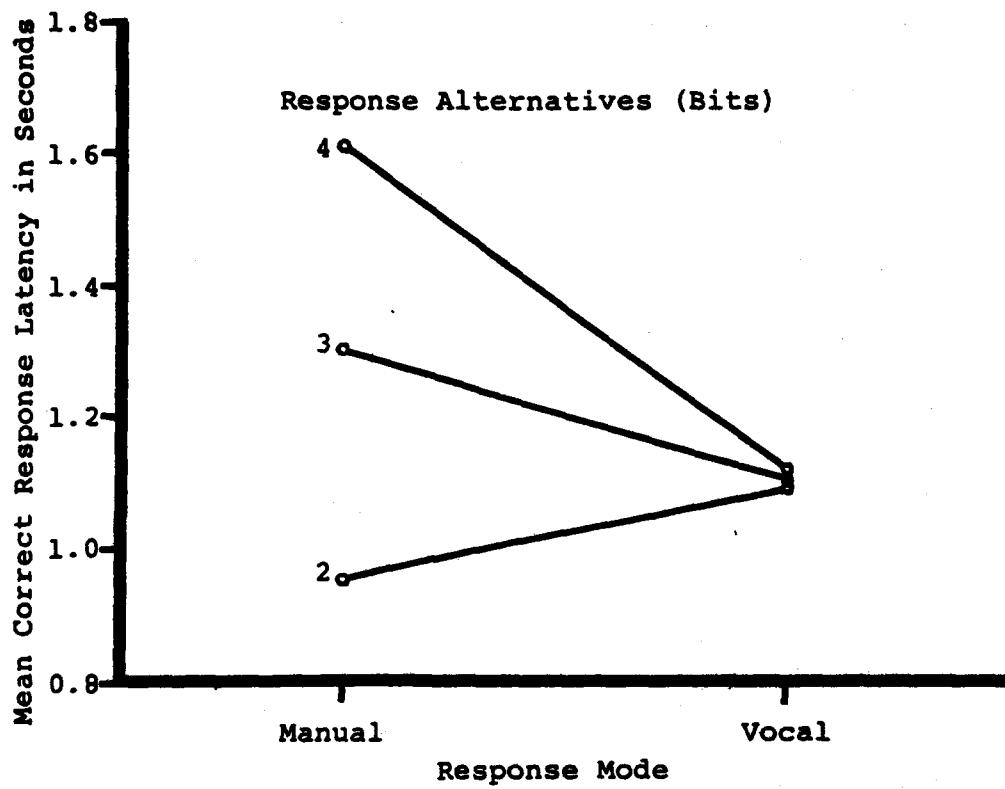
Figure 10. Mean correct response latency as a function of response mode and response alternatives in digit entry task.

subjects were not able to maintain their single task perfor-
mance level in any dual task treatment combination. The
single task means of 1.035 and 1.083 seconds at the manual
and vocal response mode are respectively superior to the
dual task means of 1.540 and 1.140 seconds. This function
is depicted in Figure 11. Figure 12 shows the same
relationship at the three levels of response alternatives
with single-task means of .949, 1.064, and 1.163 seconds,
being contrasted to 1.096, 1.345, and 1.580 seconds for the
dual-task load conditions.

The F test indicated that there were significant dif-
ferences between the levels of response alternatives in the
manual response mode for means of .954, 1.296, and 1.613
seconds at two, three, and four bits of information. The
Tukey HSD statistic confirmed that each of the means were
significantly different from each other beyond the one
percent level of significance. The response alternative
means at the vocal response mode level, however, were not
found to differ from each other. The vocal task performance
did not vary with increasing levels of response
alternatives, while the manual performance decreased dra-
matically. This relationship is graphically demonstrated
in Figure 13. A similar relationship is shown for response
alternatives at both the single and dual task load condi-
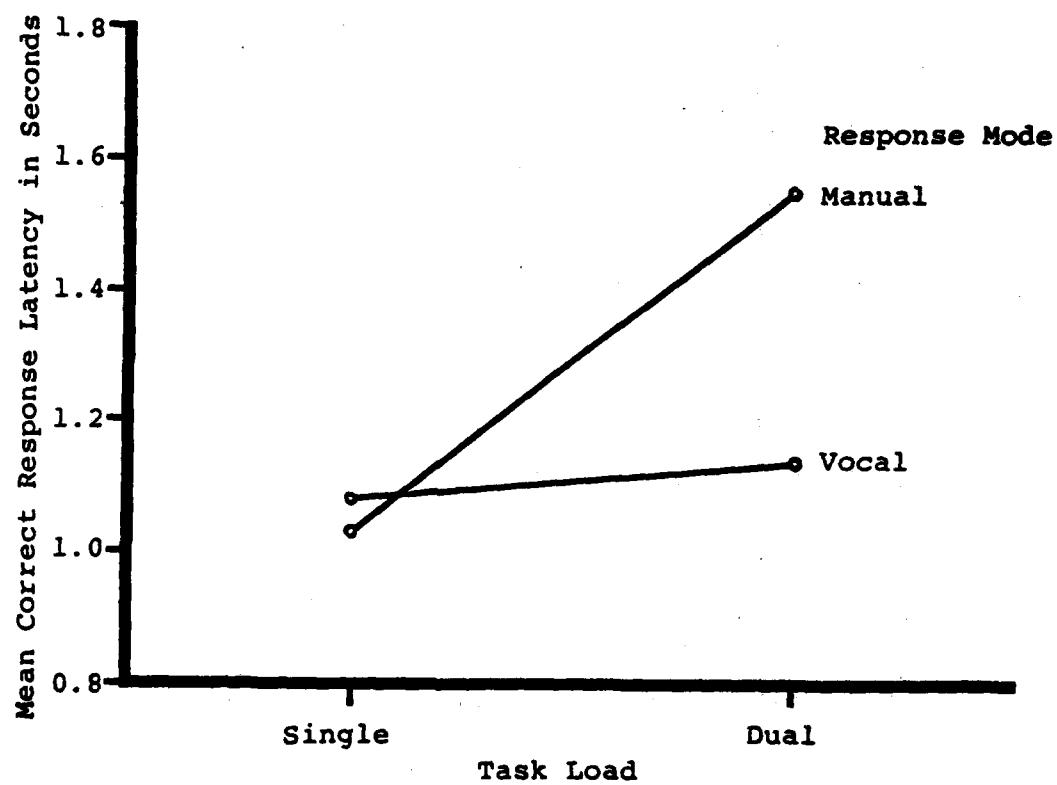
Figure 11. Mean correct response latency as a function of
task load and response mode in digit entry task.
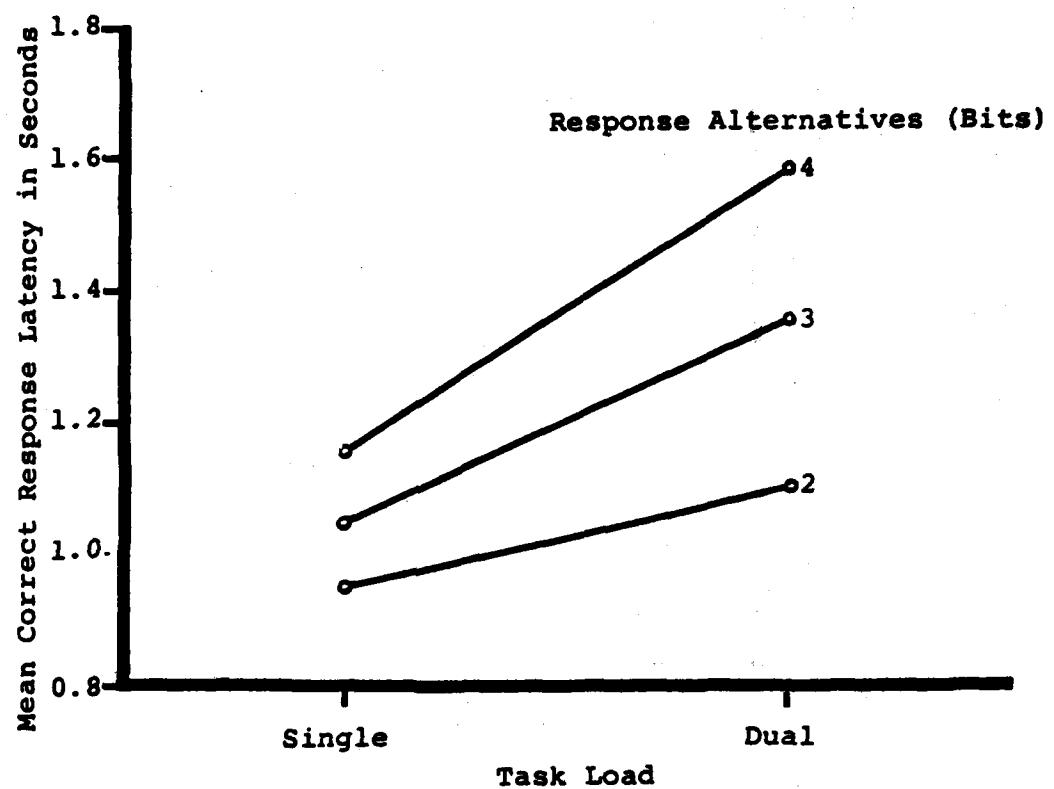
Figure 12. Mean correct response latency as a function of task load and response alternatives in digit entry task.
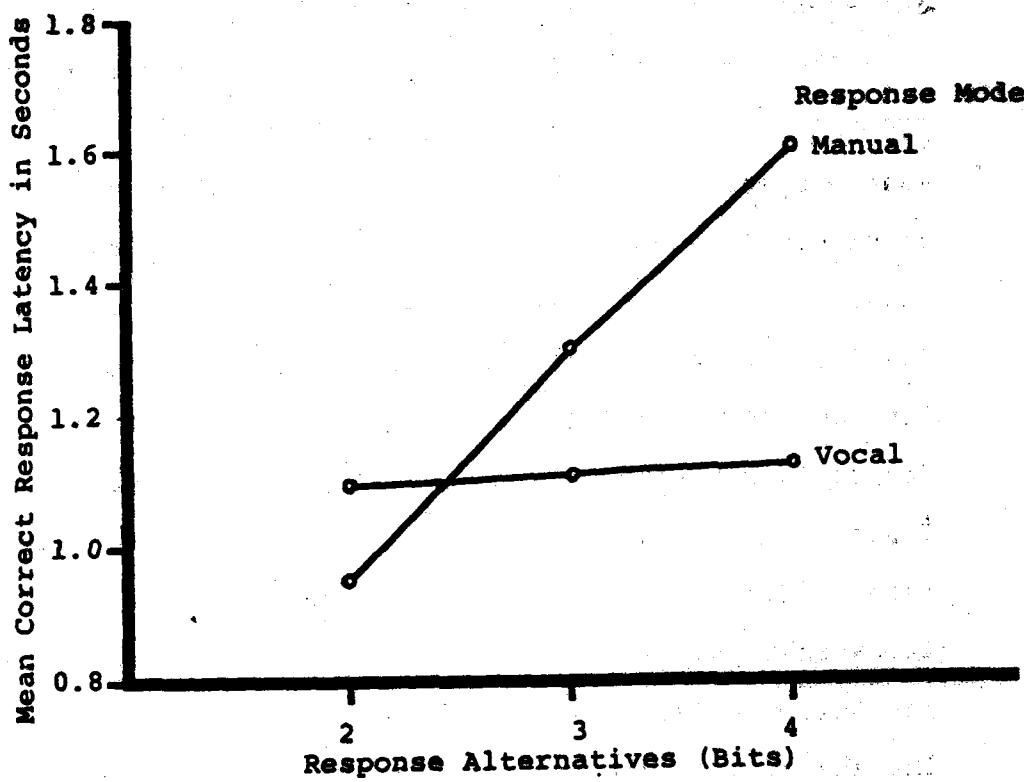
Figure 13. Mean correct response latency as a function of response alternatives and response mode in digit entry task.

tions in Figure 14. Tuckey's HSD test indicated that the response alternative means of .949, 1.064, and 1.163 seconds in the single condition and 1.096, 1.345, and 1.580 seconds latency for the dual condition were each significantly different from each other at each task load level.

The accuracy of digit entry performance was also recorded with the above latency variable. Error data for each subject was converted to an error rate score for each condition, because the number of responses emitted by each subject under the dual-task condition was not constant. The dual-task trials were self-paced, but the tracking task limited the duration of the trials. Descriptive statistics for the digit entry task response error rate may be found in Table 4. This table suggests that there are significant main effects for all three independent variables under investigation. The analysis of variance procedure employed in this study, however, does not confirm this suggestion. The analysis of variance source table for the response error rate variable is presented in Table 5.

A test of homogeneity of variance rejected the hypothesis of homogeneity of variance for the response error rate as a percentage score. Hartley's Fmax statistic (Kirk, 1968) of 15.54 exceeded the tabled value of Fmax .01/12, 16 = 8.00, therefore, the hypothesis of homogeneity was rejected.
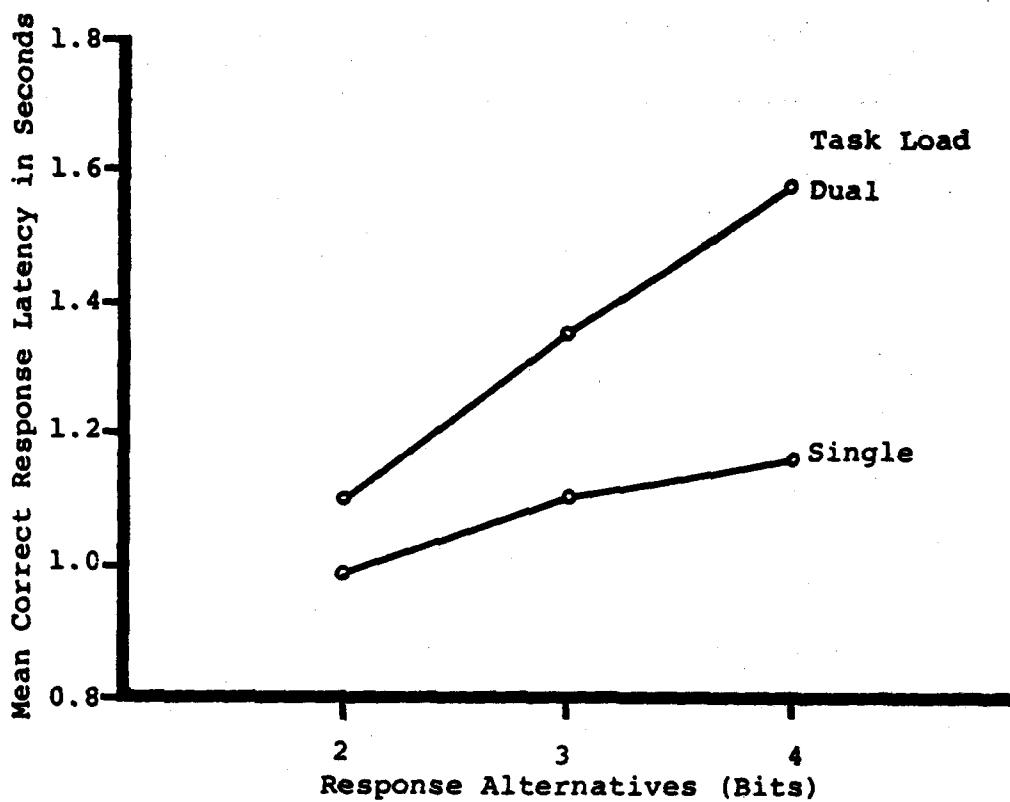
Figure 14.   Mean correct response latency as a function of
response alternatives and task load in digit
entry task.

TABLE 4

DESCRIPTIVE STATISTICS FOR RESPONSE ERROR
RATE IN DIGIT ENTRY TASK

Condition	Mean	Standard Deviation	Variance
A - Response Mode			
Manual	2.439	3.733	13.93
Vocal	.6116	1.102	1.215
B - Task Load			
Single	.9361	1.591	2.530
Dual	2.239	3.823	16.62
C - Response Alternatives			
Four	1.481	2.126	4.518
Eight	1.457	3.281	10.77
Sixteen	1.810	3.436	11.81

TABLE 5

ANALYSIS OF VARIANCE SOURCE TABLE

FOR RESPONSE ERROR RATE

Source	SS*	df	MS*	F	
Blocks	237.9	16	14.87	2.12	**
Treatments					
A - Response Mode	189.7	1	189.7	27.09	**
B - Task Load	87.06	1	87.06	12.43	**
C - Response Alt.	5.778	2	2.889	.41	NS
AB	57.75	1	57.75	8.25	**
AC	6.981	2	3.493	.50	NS
BC	11.29	2	5.643	.81	NS
ABC	4.279	2	2.139	.31	NS
Residual	1232.2	176	7.001		
Total	1832.9	203			

* $X\ 10^{-4}$

** $p < .01$

The inverse sine transformation was selected to achieve homogeneity of error variance. This transform has been found useful when the raw score is expressed as a proportion and the means and variances are proportional.

The main effects of response mode and task load were found to be significant, as was the response mode by task load interaction. The analysis of variance source table for this randomized block factorial design is presented in Table 5. The analysis of variance indicated that there was a statistically significant difference beyond the .01 level between the vocal and manual response modes with means of .61 vs 2.24 percent error rate respectively, F (1, 176) = 27.09, $p <$ .01. A reliable difference between the single and dual (.94 vs 2.2 percent) task load error rates was found with F (1, 176)= 12.43, $p <$ .01. The analysis also revealed a significant interaction effect of response mode by task load with F (1, 176)= 8.25, $p <$ .01. The analysis of variance failed to reveal any statistically significant differences for the response alternative variable alone or in interaction with the other two variables.

The analysis of variance for the simple main effects of response mode by task load are presented in Table 6. All the statistics were found to be statistically significant

TABLE 6

ANALYSIS OF VARIANCE TABLE FOR SIMPLE MAIN EFFECTS

OF RESPONSE ERROR RATE

Source	SS	df	MS	F
A - Response Mode				
A at b1	4.776	1	4.776	6821.5 *
A at b2	149.7	1	149.7	213807 *
B - Task Load				
B at a1	148.1	1	148.1	211478 *
B at a2	4.488	1	4.488	6409.9 *
Residual	.1232	176	.0007	

* $p < .01$

beyond the .01 level. Figure 15 shows mean error response rate as a function of both the response mode and task load variables. The manual mode was found to have a higher error rate than the vocal mode at both the single and the dual task load levels with means of 1.37 vs .504 and 3.73 vs .746 respectively. The dual-task load was also found to have significantly higher error rate for both the manual and vocal response mode conditions with means of 3.73 vs 1.37 and .746 vs .504 percent respectively.

The number of times that the voice recognition unit failed to correctly identify a vocal response, when the subject did respond correctly, was recorded. The machine error rate for each condition was calculated for each subject from this score in the same manner in which the response mode error rate was derived. A graphic comparison of the machine error rate and the subject error rate indicated that the machine error rate was an order of magnitude greater than the human error rate with means of 17.94 vs 1.587 percent. See Table 7 for further descriptive statistics of the machine error rate.

A randomized block factorial design--RBF-pq (Kirk, 1968)--was selected to test the effects of the task load and response alternative variables upon the machine error rate. The analysis of variance source table is presented in
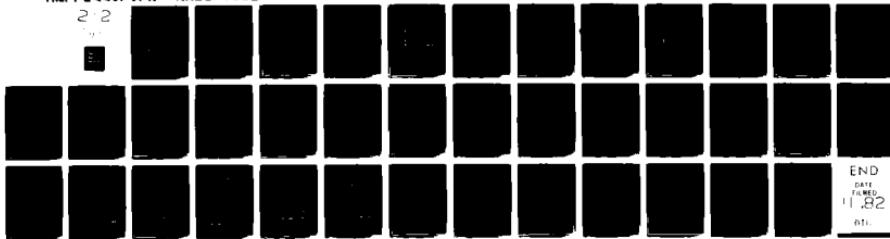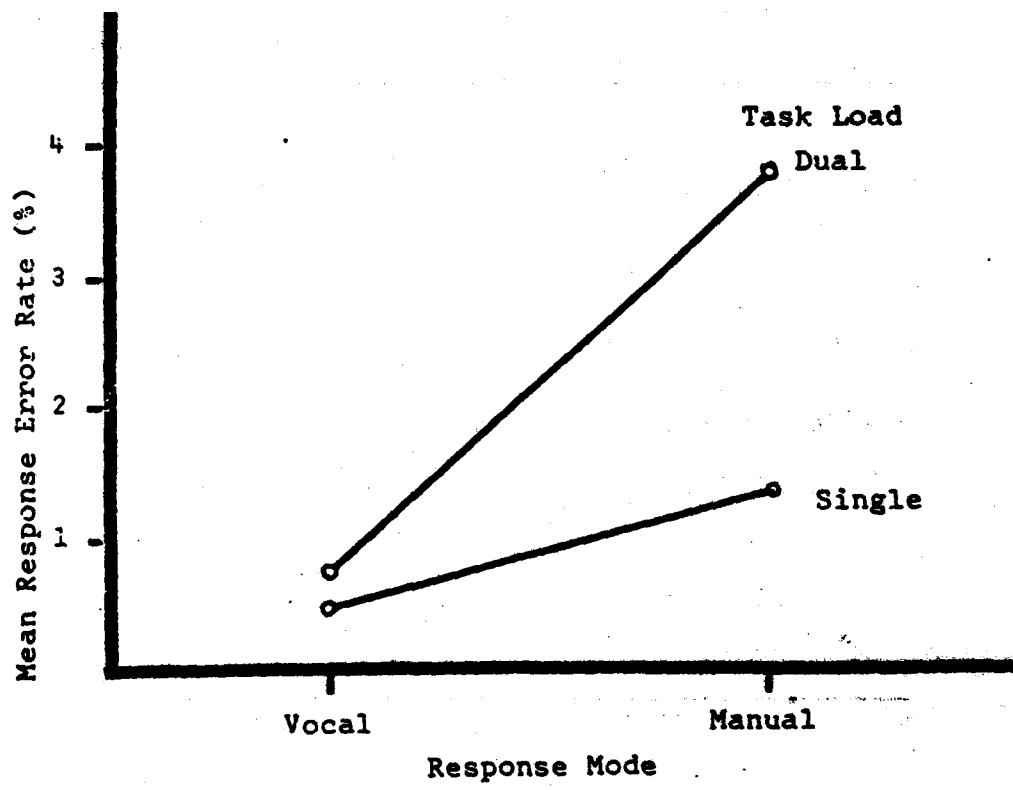
Figure 15. Mean response error rate as a function of response mode and task load in digit entry task.

TABLE 7

DESCRIPTIVE STATISTICS FOR MACHINE RECOGNITION
ERROR RATE IN DIGIT ENTRY TASK

Condition	Mean	Standard Deviation	Variance
A - Task Load			
Single	17.81	11.26	126.8
Dual	17.96	12.75	162.5
B - Response Alternatives			
Four	16.06	12.94	167.6
Eight	17.81	13.50	182.2
Sixteen	19.96	9.324	86.93

Table 8. The Fmax statistic would have to exceed Fmax .05/6, 16 = 4.37 in order to reject the hypothesis of homogeneity of variance. Since Fmax was equal to 3.1, the hypothesis was not rejected and no further transformation was considered. There were no statistically significant differences in machine error rate associated with the two independent variables. Figure 16 only suggests a relationship between machine error rate and the variables of task load and response alternatives.

Tracking Task

The dependent variable in the tracking task was the root mean square error (RMSE) in percent of tracking scale. The descriptive statistics for this variable are tabulated in Table 9. An examination of this table will reveal that the RMSE score varies as a function of the task load but remains constant at different levels of the response alternatives. An analysis of variance source table is presented in Table 10. It also indicates that the variable of task load had a significant effect upon tracking performance.

A randomized block factorial design (designated RBF-pq by Kirk, 1968) was selected as the analysis of variance procedure to evaluate the effects of three levels of task load and three levels of response alternatives upon tracking

TABLE 8

ANALYSIS OF VARIANCE SOURCE TABLE FOR MACHINE
RECOGNITION ERROR RATE IN DIGIT ENTRY TASK

Source	SS	df	MS	F	
Blocks	7529.5	16	470.6	7.508	*
Treatments					
A – Task Load	.0460	1	.0460	.00073	NS
B – Response Alt.	257.53	2	128.8	2.054	NS
AB	97.80	2	48.90	.7802	NS
Residual	5014.4	80			
Total	12899	101			

* $p < .01$
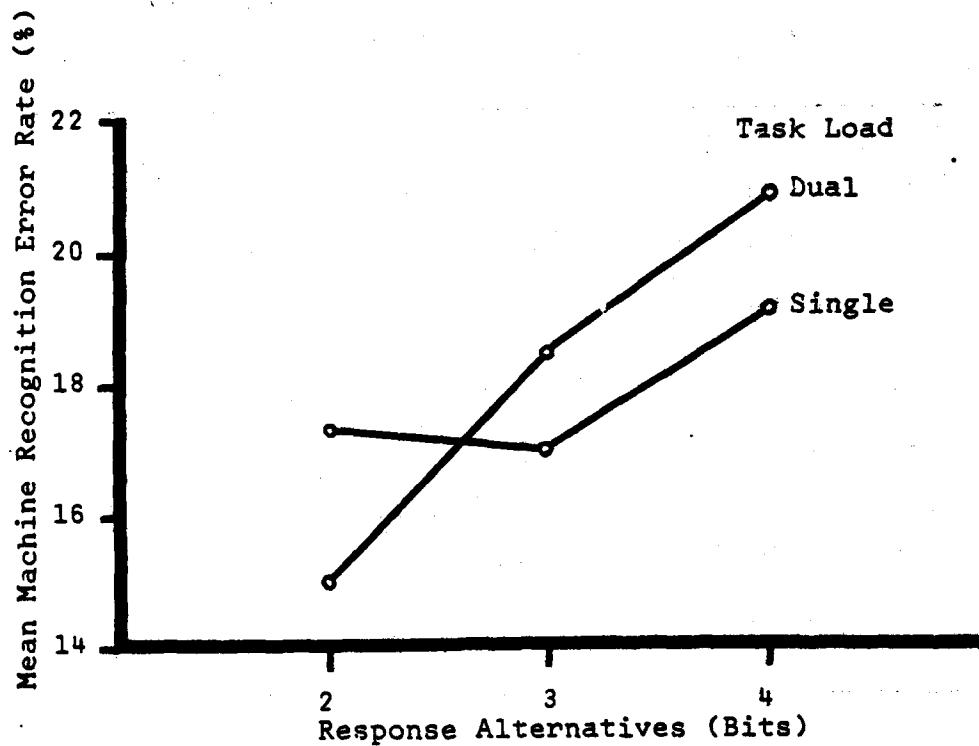
Figure 16. Mean machine recognition error rate as a function of response alternatives and task load in digit entry task.

TABLE 9

DESCRIPTIVE STATISTICS FOR RMS ERROR

IN TRACKING TASK

Condition	Mean	Standard Deviation	Variance
A - Task Load			
Tracking Single	.163	.0637	.0041
Manual Dual	.406	.0896	.0080
Vocal Dual	.216	.0865	.0075
B - Response Alternatives			
Four	.262	.1308	.0171
Eight	.263	.1381	.0191
Sixteen	.263	.1291	.0167

TABLE 10

ANALYSIS OF VARIANCE SOURCE TABLE FOR RMS

ERROR IN TRACKING TASK

Source	SS	df	MS	F
Blocks	.6490	16	.0406	16.748**
Treatments	1.9941			
A - Task Load	1.6720	2	.8360	345.197**
B - Response Alt.	.0002	2	.0001	.052 NS
AB	.0119	4	.0029	1.239 NS
Residual	.3099	128		
Total	2.644	152		

** $p < .01$

RMSE. An Fmax statistic was computed for the nine treatment combinations. A value of Fmax $(9, 16) = 2.895$ was not great enough to reject the hypothesis of the homogeneity of variance.

The main effect of task load was significant with F $(2, 128) = 345.197$, $p < .01$. The response alternative main effect was not found to be significant, nor was the simple main effect of task load by response alternatives. The above relationship is illustrated in Figure 17, in which RMS error is plotted as a function of task load. Figure 18 illustrates this relationship for the response alternatives as well.

The Tukey HSD statistic revealed that the means of .162, .216, and .406 percent scale for the tracking task performed singly, with the vocal task, or the manual digit entry task were all found to differ from each other with a significance level of $p < .01$.
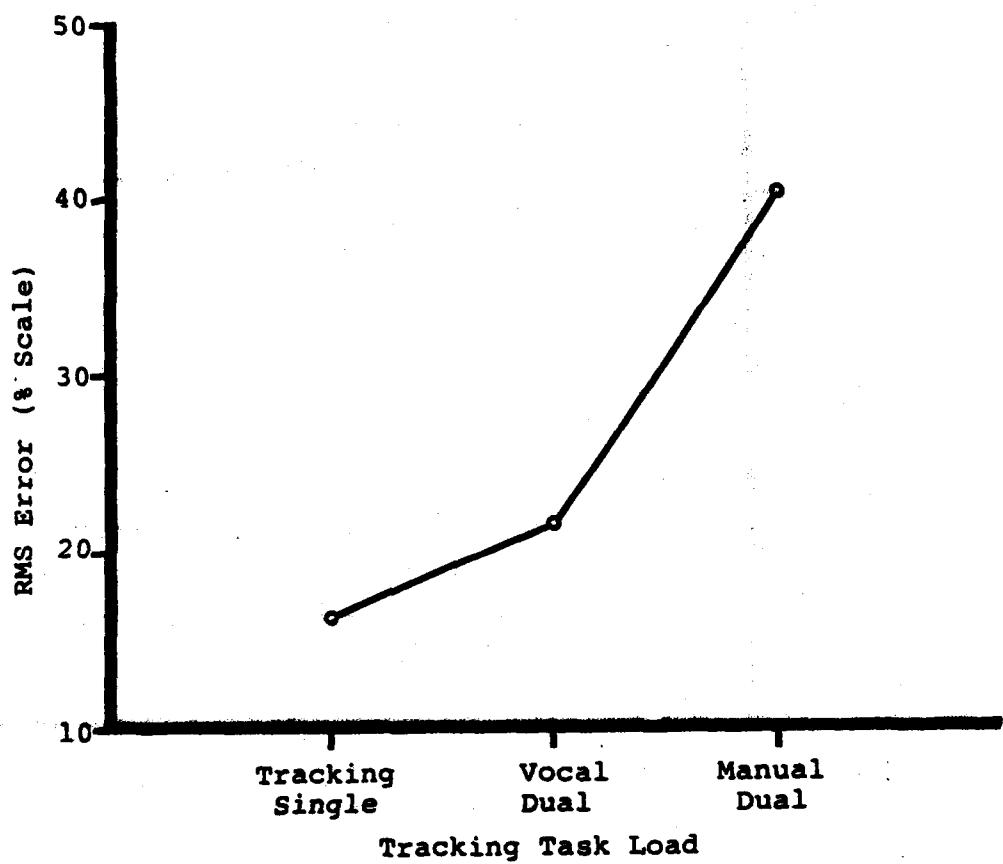
Figure 17. Tracking RMS error as a function of task load in tracking task.
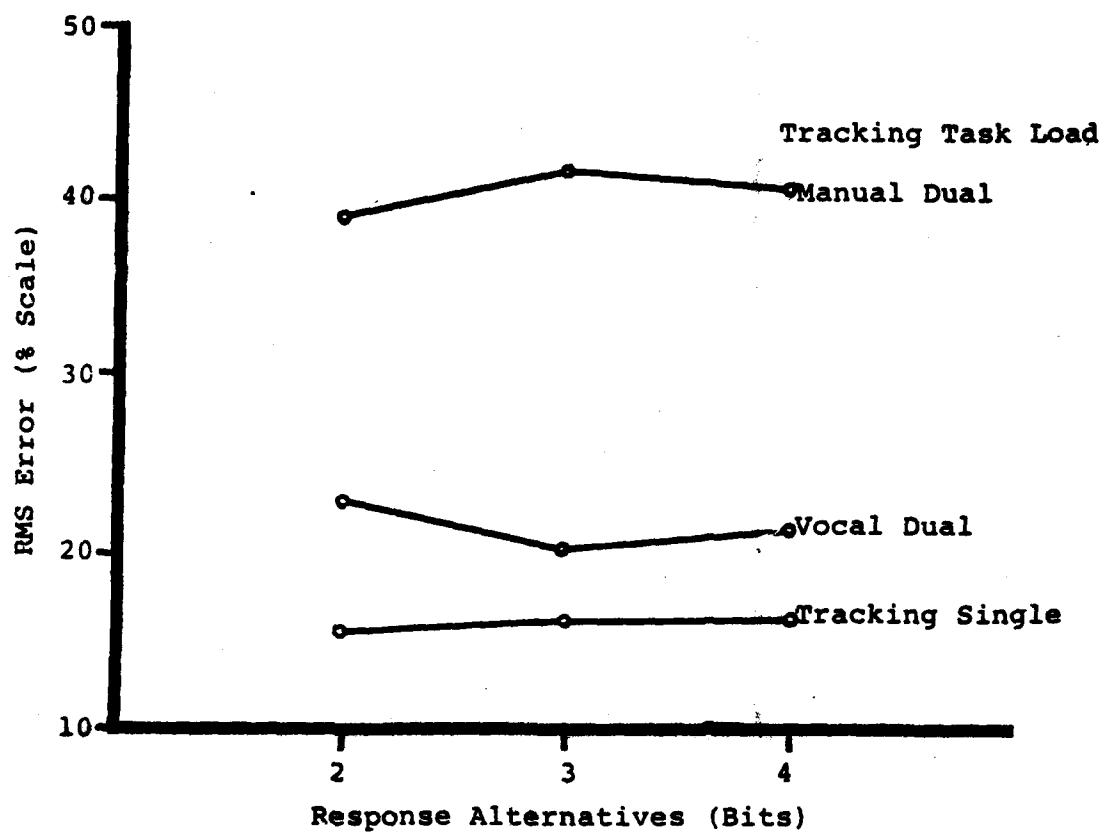
Figure 18.   Tracking RMS error as a function of response
alternatives and task load in tracking task.

CHAPTER IV

DISCUSSION

The present research was designed to test several assumptions regarding the influence of response mode, response alternatives, and task load on complex multi-task performance.

Proponents of the application of voice recognition technology to man-machine systems (Hill, 1972; Note 2), have suggested that the vocal response mode superior to the manual response mode in several aspects. Results from the present study demonstrate that the vocal response mode is dramatically more effective in terms of speed and accuracy than manual mode when it comes to high task difficultty in a time-sharing situation. The hypothesis that the vocal response mode is faster than the manual response mode was rejected in favor of a more general alternative hypothesis that speed of response mode is a function of task load and the number of response alternatives.

The analysis of variance in Table 2 indicated that the main-effect of vocal response mode had a correct response latency mean that was .18 seconds lower than the manual response mode. While this difference was significant at the

97

$p<$ .01 level, it obscured the fact that at some treatment combinations the difference was much greater, and that at other treatment combinations the direction of the difference was reversed. The tests of simple main effects (Table 3) for response mode at single and dual-task loads and four, eight, and sixteen response alternative treatment levels were all significant beyond the $p<$ .01 level. Figures 9 and 11 illustrate that the manual response mode was slightly faster than the vocal mode in the single-task condition by .05 seconds. The vocal response mode, however, was notably superior to the manual mode under dual-task conditions, with a difference of .40 seconds. The manual response mode was again faster than the vocal mode at the four response alternative level by .14 seconds. The vocal response mode had a lower latency at the eight and sixteen response alternative treatment levels by a greater margin of .18 and .48 second respectively. These differences are depicted in Figures 10 and 13.

In the Braunstein and Anderson study (Note 4) the subjects in a digit entry task could read the digits faster than they could keypunch, even after several hours of practice. Comparatively higher digit input rates were reported in the above experiment for both response modes, possibly because the digits were presented in a four-digit

numeral. Furthermore, the subjects had several hours of practice and there was no requirement for a disciplined voice response.

It can be seen from the present study that the manual response mode is marginally faster at both the single task conditions and the four response alternative conditons, while the margin of superiority was much better for the vocal response mode at the more complex treatment combinations. This effect corroborates the findings of Welch, (Note 5), who found the vocal mode faster for two complex conditions of text entry and alpha-numeric string entry in a dual-task situation. The manual response mode was faster for the entry of alpha-numeric data of three and ten characters, utilizing a standard typewriter keyboard. It should be noted that the subjects who were inexperienced with the typewriter keyboard performed better with the voice recognition device in the entry of complex data. A similar familiarity effect might be contributing to the speed of vocal entry in the present study. It is highly probable that several more hours of practice with the keyboard used in this study would have produced better manual performance.

The single-channel hypothesis (Broadbent, 1958; Welford, 1971; and others) implies that as the information load increases, it will take longer to process each item.

This suggests that the increasing response latencies will be a function of increasing response alternatives. This hypothesis is clearly supported in the Hick (1952), Hyman (1953), and Damos and Wickens (1977) studies as reported earlier. The results from this study as illustrated in Figure 13 clearly show this relationship. The correct response latencies do not, however, increase with increasing numbers of response alternatives for the vocal response mode. The slope of this function is flat and identical to reported results in the studies of Alluisi and Muller (1958); Mowbray and Rhoades (1959); Mowbray (1960); and Leonard (1959), which suggested that the performance results were due to over-learning, extended practice, or high compatibility.

A very strong possibility is that the manual performance decrement is due to what Kahneman (1973) has identified as structural interference. If a subject had not memorized the keyboard, the correct key would have to be searched for and presssed. If this is the case, the single-channel hypothesis could not be supported with the manual mode results from the present study. In any case, Hill's (1972) suggestion that the vocal response mode is independent of factors that normally affect sight and reach is suppported by the above results.

The single-channel concept also suggests that time-sharing the operator's resources by performing the digit digit entry task simultaneously with the tracking task, will produce interference with the performance of one or both of the tasks with respect to each of the task's isolated performances. The analysis of simple main effects in Table 3 shows that the latency increased from the single-task condition to the dual-task performance for both the manual (148.80%) and vocal (105.26%) response modes which were statistically significant at $p<.01$. This relationship is illustated in Figures 9 and 11. It is possible that the subjects might have protected the performance level of the harder task (Mowbray, 1953; Kahneman, 1973) by trading off performance on the easier task. The tracking task was generally considered the hardest of the tasks. The error score of the tracking task increased 249.08 percent, when tracking was paired with the manual digit entry task, and 132.52 percent, when it was paired with the vocal entry task. The analysis of variance for the main-effect of task load indicated that these differences were statistically significant with $p<.01$. These differences are illustrated in Figure 15. If any task performance was protected, it must have been the digit entry task. The above results indicate that the vocal response mode does interfere less

with other activities than the manual mode does. (Hill, 1972). The results also suggest that either 1) the subjects could not manage the task of giving equal priority to each of the dual-tasks; or 2) that the feedback display did not provide the subjects with sufficient information to manage equal performance levels; or 3) the combination of the manual entry task and the tracking task was so difficult that the subjects ignored the feedback display.

While it is possible that a proportion of the mutual interference effects of the manual entry task and the tracking task are structural in nature, there is little evidence to suggest that there is anything but central interference occuring in the dual vocal-tracking task.

The single-channel hypothesis would have been further supported if the main effect of response alternatives had been statistically significant in the tracking task. Damos and Wickens (1977) reported that tracking performance in a similar manual dual-task situation deteriorated with increasing levels of response alternatives. They attributed this to structural interference, i.e. looking away from the tracking display in order to find a key. The results from the present study indicate, however, that there was no additional interference with the tracking performance due to visually searching for additional keys. The Damos and

Wickens study did not utilize a feedback display. A question for further research is whether or not the feedback enabled the subjects to trade-off performance at the expense of the digit entry task. Figure 14 suggests that this might be the case in the present study.

A somewhat unexpected finding of this experiment was the extremely high error rate generated by the voice recognition unit. Machine "mis-recognition" errors were more than ten times greater than the human errors. Given situations of multiple functions and high task loads, the performance of the vocal system--human and machine combined--is obviously completely unacceptable.

In this experiment the voice recognition machine errors were recorded separately from the subject errors, it can be seen that the vocal data entry by itself has less errors than manual entry and that the vocal response mode was the most error-free condition. The Harris, et al (Note 8) study had reported that vocal digit entry had lower accuracy than manual keyboard entry. But since their vocal data included errors caused by machine "mis-recognition" of the spoken word, these results are not directly comparable.

It is also of interest to note that only the variables of response mode and task load had any significant effect on the human error rate, while apparently response alternatives

had no effect on accuracy. In other words, the error rate stayed the same regardless of the fact that the subject had four, eight or sixteen choices. Looking at this result from a single-channel viewpoint, an increase in response alternatives should have been followed by an increased error rate, but the data do not support this assumption.

Welch (Note 5) found in a serial digit entry task that the manual mode had less errors than the vocal input mode. When the entry task included alphabetic characters as well, the vocal errors were lower than that of the manual mode. The data is not directly comparable, but with the addition of a time-shared task, the Welch study also found both speed and accuracy to be superior in the vocal mode.

It is to be expected that untrained subjects performed more accurately verbally than they did manually. Braunstein and Anderson (Note 4) reported in a single-task study with untrained operators that for numeric entry the vocal error rate was equal to the manual error rate, while Alluisi and Muller (1958) found that the vocal error rate was lower than that of the manual mode.

There are several implications that can be made from this study that might not have been apparent from previous investigations. At the present state-of-the-art in voice recognition systems with isolated word recognition capa-

bilities, there is little reason to consider incorporating such equipment unless there are more than four functions to be performed and vision is occupied elsewhere. Also, at the present time, the voice recognition device can not accept input at the rate that the operator is capable of speaking. If 1/4 second could be subtracted from each of the vocal response latencies, then the vocal response mode would be faster than the manual mode at all treatment levels. If the voice recognition hardware could be developed to recognize a spoken utterance 1/4 of a second faster, the human operator would be able to take advantage of this improvement. Harris, North and Owens (Note 8) reported that the recognition unit employed in the present experiment required a mean of .531 seconds to identify the digits one through eight. This did not include the time for the duration of the word. An improvement in voice recognition speed would certainly lead to improvements in total system performance.

The reliability and utility of the two dependent variables in the digit entry task could be improved by minor modifications in the voice recognition unit and the performance recording technique. The speed and accuracy measurements were made in such a manner that it was impossible to automatically discriminate between the reac-

tion time of the subject and the recognition processing time of the machine. The errors made by the human operator and the errors of the voice recognition unit were recorded by the experimenter. It is recommended that follow-up research employ a voice recognition device that can measure reaction time from the onset of the stimulus to the onset of the vocal response. Since a considerable number of researchers report this measurement, it would be very useful for comparison purposes. The duration of the response word might also yield another dependent variable that might be sensitive to changes in work load and stress. The machine processing time is also partially a function of how well the operator can repeat the response words consistently.

The results from this study do suggest the possibility that the machine error rate increased as a function of increasing response alternatives. This could prove to be an interesting index of work load or stress associated with task difficulty. A measure of recognition reliability was computed internally in the voice recognition device, but it was not available for observation. This index could also serve the same function as the word duration variance in estimating work load, while the subject is concurrently performing another task. Another source of potential errors could be eliminated if both the vocalization and the machine

recognition parameters could be recorded for later analysis. The weakest link in the recording of the error data in this experiment was that the experimenter had to monitor both the stimulus list and the subject for response errors.

It is highly recommended that any future research corroborate the findings with respect to the manual digit entry performance with a different keyboard. It is possible that manual reaction time could be reduced with a lighter activation force keyboard. Two other keyboards that might be of interest in future studies of this kind, are the chord keyboard and the multi-function keyboard. While the chord keyboard requires extensive training, the advantage of not having to visually search for the keys would rule out the effect of structural interference confounding the response data. The multi-function keyboard has a limited number of switches with a larger number of changeable legends. This keyboard is gaining in acceptance as designers are running out of panel space in many crew-stations. As with the chord keyboard, the multi-function keyboard has the advantage of reducing finger movement and search time. Both these keyboards can provide a higher number of response alternatives, which would improve on the possibility of a fair comparison with vocal control.

In several specific instances in this experiment, some

of the subjects could have achieved a better RMS error tracking score if they had left the hand-control alone. It is suggested that a more sensitive parameter than RMS error could be found that would correlate positively with increasing response alternatives.

A performance feedback display that does not load the subject down with an additional task of its own, would be a valuable aid in evaluating the results of any future research effort in this area.

Any future attempts to fairly evaluate the effectiveness of the vocal response mode should also incorporate more time for the training of subjects in learning how to interface with the voice recognition unit. A method of verifying that the machine correctly recognized the subject's response, would also reduce the number of machine errors and increase the man-machine system performance. Perhaps a speeded synthesized voice response unit might successfully serve this function.

There are several recommendations that can be made to system designers of such systems as the advanced airborne weapons system. However, the data do not support at this time the vocal activation of any type of weapons. The level of reliability of the man-machine system is unacceptable for such a critical task. With the development of reliable

verification and correction techniques, vocal input can become a viable alternative to many manual systems that 1) require more than six functions to activate; 2) require a faster response time than that possible with the manual entry; 3) require better accuracy than that of the manual; and 4) find the reliability of the machine recognition to be acceptable.

A recommendation can be made for systems with many functions, low sensitivity to errors, and high turnover in personnel. The high cost of the voice recognition equipment can be offset by reduced training time and reduced skill requirements of operator personnel.

The dual-task testing procedures used in this study could prove to be useful in evaluating the effects of more complex data entry tasks or the effects of acceleration and vibration on data entry performance.

A final implication is that at the present time the number of manual switching functions which can be implemented in a crew-station is limited by a function of the available panel space and, according to this study, the interation of task load and number of response alternatives. There is clearly no such limit with the vocal response mode with voice recognition equipment, other than the hardware limitations of computer speed and memory size.

SUMMARY AND CONCLUSIONS

In the increasingly complex area of man-machine interface, the human operator is often required to perform two or more tasks which demand full attention simultaneously, oftentimes overloading the operator's manual and visual capabilities. The recent commercial availability of voice recognition devices has opened up the possibility that a voice-actuated response could act as an extra pair of hands for the operator.

Proponents of voice recognition devices have hypothesized that vocal response would be independent of other channels; would interfere less with other activities and would be faster than manual response mode. The present study was designed to test the effectiveness of voice response as compared with a manual response in a time-sharing environment.

The subjects were seventeen paid volunteers from a subject pool available to the Naval Aerospace Medical Research Laboratory in Pensacola, Florida. All subjects performed a digit entry task concurrently with a one-dimensional tracking task. Subjects performed all three tasks

isolation. The digit entry task required the subjects to vocalize a digit or press a key corresponding to a digit presented on a CRT display. Vocal input was recorded via a voice recognition device. Manual input was through a sixteen key keyboard. The compensatory tracking task required appropriate left-right movements of a handcontrol to maintain the position of a diamond-shaped cursor on a vertical line in the center of a horizontal track. This was also presented on the CRT. In addition, a feedback indicator presented the subjects knowledge of their performance on any one of their current tasks. Following a training period, the subjects performed a series of fifteen three minute trials in which their performance was tested on the three tasks in isolation. In addition, the manual task and the vocal task were performed concurrently with the tracking task.

A randomized block factorial design was used in evaluating the effects of three levels of response alternatives (four, eight, or sixteen digits in a set) and three levels of tracking task load (tracking alone, tracking with manual entry, or tracking with vocal entry) upon tracking performance. A randomized block factorial design was also used in evaluating the effects of three levels of response alternatives, two levels of response mode and two

levels of task load upon the performance of a digit entry task. The response alternatives were four, eight, or sixteen digit sets selected from the digits one through sixteen; response mode was the manual or vocal entry of the digits; task load was the performance of the task by itself or concurrent with the tracking task. The subjects were assigned a random order of block treatments. The dependent variables recorded in this study were: 1) digit entry response latency interval measured from start of digit presentation to the registration of a response to that digit; 2) digit entry errors, and 3) tracking root mean square error.

Several of the major hypotheses tested supported the single-channel concept. The results indicated that manual entry performance declined with increased number of response alternatives, and that both manual and vocal performances deteriorated with increased task load, although deterioration was significantly less with the vocal input mode. Tracking performance also decreased with dual-task load. Response alternative loading had no significant effect on the performance of the vocal or tracking task. The manual mode was found to have a higher error rate than the vocal mode at both the single and the dual task load levels. The dual-task load was also found to have a

significantly higher error rate for both the manual and vocal response mode conditions.

Findings unique to this study implied that the point at which vocal performance became equal to or better than the manual performance was at the eight response alternative level in a single task situation and at four alternatives in a dual-task situation. Dual-task performance also caused interference for both time-shared tasks which could not be controlled by providing the operator with a feedback display.

Results from this study demonstrate that the vocal response mode is dramatically more effective in terms of speed and accuracy than manual mode when it comes to high task difficulty in a time-sharing situation. The hypothesis that the vocal response mode is faster than the manual response mode was rejected in favor of a more general alternative hypothesis that speed of response mode is a function of task load and the number of response alternatives.

Reference Notes

1. Lane, N. E., Strieb, M. I., & Wherry, Jr., R. J.   The human operator simulator: Workload estimation using a simulated secondary task. Paper presented at the NATO/AGARD Aerospace Medical Panel Specialists Meeting, Koln, Germany, April, 1977.

2. Feuge, R. L. A technical review of government applicacations of automatic speech technology (AST). Prepared for the Office of Naval research and Air Development Center under contract No. N00014-77-C-0141. September 1977.

3. Navon, D. & Gopher, D. On the economy of the human processing system: A model of multiple capacity. Unpublished manuscript, The Technion-Israel Institute of Technology, Haifa, Israel, 1977.

4. Braunstein, M. & Anderson, N. S. A comparison of reading digits aloud and keypunching. IBM Research Center, (Research Memorandum RC-185). Yorktown Heights, N. Y., November 1959.

5. Welch, J. R. Automatic data entry devices. (Tech. Rep. Contract No. F 30602-76-C-0178). Air Force Systems Command Base, New York, 1977.

6. Galanter, E., & Owens, J. A. The independence of the simple reaction time on key force (Tech. Rep. PLR-31). New York: Columbia University, Psychophysics Laboratory, July, 1974.

7. Gopher, D. & North, R. A. The measurement of operator capacity by manipulation of dual-task demands. Savoy, Ill.: University of Illinois at Urbana-Champaign, Institute of Aviation, Aviation research Laboratory, Technical Report ARL-74-21/AFOSR-74-15, 1974.

8. Harris, S. D., North, R. A., & Owens, J. M. A system for the assessment of human performance in concurrent verbal and manual control tasks. Paper presented at the seventh annual meeting of the National Conference on the use of on-line computers in psychology, Washington, D. C., November, 1977.

Reference List

Alluisi, E. A. Information and uncertainty: The metrics of communications. In K. B. DeGreene (ed.), *Systems psychology*. New York: McGraw Hill, 1970.

Alluisi, E. A. & Miller, P. F., Jr. Verbal and motor responses to seven symbolic visual codes: A study in S-R compatability. *Journal of Experimental Psychology*, 1958, 55, 247-254.

Beek, B., Neuberg, E. P., & Hodge, D. C. An assessment of the technology of automatic speech recognition for military applications. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 1977, ASSP-25 (4), 310-319.

Brainard, R. W., Irby, T. S., Fitts, P. M., and Alluisi, E. A. Some variables influencing the rate of gain of information. *Journal of Experimental Psychology*, 1962, 63 (2), 105-110.

Broadbent, D. E. *Perception and Communication*. London: Pergamon Press, 1958.

Broadbent, D. E. *Decision and Stress*. New York: Academic Press, 1971.

Chapanis, A. Psychology and the instrument panel. *Scientific American*, 1953, 188:18, 76-82.

Cliff, R. C. *Attention sharing in a dynamic multi-task environment*. Unpublished doctoral dissertation, Dept. of Industrial Engineering and Operations Research, University of California, Berkeley, September 1971.

Cliff, R. C. Attention sharing in the performance of a dynamic dual task. *IEEE Transactions on Systems, Man, and Cybernetics*, 1973, SMC-3, (3), 241-248.

Damos, D. L. & Wickens, C. D. Dual-task performance and the Hick-Hyman law of choice reaction time. *Journal of Motor Behavior*, 1977, 9, (3), 209-215.

Deutsch, J. A. & Deutsch, D. Attention: Some theoretical considerations. *Psychological Review*, 1963, 70, 80-90.

Fitts, P. M. Engineering psychology. In <u>Annual Review of Psychology</u>, 1958, <u>9</u>, 267-294.

Fitts, P. M. & Posner, M. I. <u>Human Performance</u>. Belmont, California: Brooks/Cole, 1967.

Fitts, P. M. & Seeger, C. M. S-R compatability: spatial characteristics of stimulus and response codes. <u>Journal of Experimental Psychology</u>, 1953, <u>46</u>, 199-210.

Fogel, L. J. <u>Human Information Processing</u>. Englewood Cliffs, N. J.: Prentice-Hall, 1967.

Gopher, D. & Kahneman, D. Individual differences in attention and the prediction of flight criteria. <u>Perceptual and Motor Skills</u>, 1971, <u>33</u>, 1335-1342.

Haber, R. N. Information processing. In C. E. Carterette & M. P. Friedman (Eds.), <u>Handbook of Perception</u> (Vol. 1), New York: Academic Press, 1974.

Hick, W. E. On the rate of gain of information. <u>Quarterly Journal of Experimental Psychology</u>, 1952, <u>4</u>, 11-26.

Hill, D. R. An abbreviated guide to planning for speech interaction with machines: the state of the art. <u>International Journal of Man-Machine Studies</u>, 1972, <u>4</u>, 383-410.

Hyman, R. Stimulus information as a determinant of reaction time. <u>Journal of Experimental Psychology</u>, 1953, <u>45</u>, 188-196.

Johnston, W. A., Greenberg, S. N., Fisher, R. P., & Martin, D. W. Divided attention: A vehicle for monitoring memory processes. <u>Journal of Experimental Psychology</u>, 1970, <u>83</u> (1), 164-171.

Kahneman, D. <u>Attention and Effort</u>. Englewood Cliffs, New Jersey: Prentice-Hall Inc., 1973.

Koster, W. G. (Ed.). <u>Attention and Performance II</u>. Amsterdam: Worth-Holland, 1969.

Leonard, J. A. Tactual choice reactions. Part I. <u>Quarterly Journal of Experimental Psychology</u>, 1959, <u>11</u>, 76-83.

Lindgren, N. Speech- Man's natural communication. IEEE Spectrum, 1967, 4 (6), 75-86.

McLeod, P. D. Interference of "attend to and learn" task with tracking. Journal of Experimental Psychology, 1973, 99 (3), 330-333.

Mowbray, G. H. Simultaneous vision and audition: the detection of elements missing from overlearned sequences. Journal of Experimental Psychology, 1952, 44, 292-300.

Mowbray, G. H. Simultaneous vision and audition: the comprehension of prose passages with varying levels of difficulty. Journal of Experimental Psychology, 1953, 46, 365-372.

Mowbray, G. H. The perception of short phrases presented simultaneously for visual and auditory reception. Quarterly Journal of Experimental Psychology, 1954, 6, 86-92.

Mowbray, G. H. Choice reaction times for skilled responses. Quarterly Journal of Experimental Psychology, 1960, 12, 193-202.

Mowbray, G. H. & Rhoades, M. V. On the reduction of choice reaction times with practice. Quaterly Journal of Experimental Psychology, 1959, 11, 16-23.

Neisser, U. Cognitive Psychology. Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1967.

Noble, M., Trumbo, D., & Fowler, F. Further evidence on secondary task interference in tracking. Journal of Experimental Psychology, 1967, 73 (1), 146-149.

Norman, D. A. Toward a theory of memory and attention. Psychological Review, 1968, 75, 522-536.

Norman, D. A. Memory and attention: An introduction to human information processing. New York: John Wiley, 1969.

Norman, D. A. & Bobrow, D. G. On data-limited and resource-limited processes. Cognitive Psychology, 1975, 7, 44-64.

Norman, D. A. & Bobrow, D. G. On the analysis of perfor-
mance operating characteristics. _Psychological Review_,
1976, _83_ (6), 508-510.

Schvaneveldt, R. W. Effects of complexity in simultaneous
reaction time tasks. _Journal of Experimental Psycho-
logy_, 1969, _81_, 289-296.

Seibel, R. Discrimination reaction time for a 1,023
alternative task. _Journal of Experimental Psycho-
logy_, 1963, _66_, 215-226.

Teichner, W. H. Interactions of behavioral and physiolog-
ical stress reaction. _Psychological Review_, 1968,
271-291.

Teichner, W. H. _Quantitative models for predicting human
visual/perceptual/motor performance_ (NMSU-ONR-TR-74-3).
Arlington, Va.: Office of Naval Research, Engineering
Psychology Programs, October 1974. (NTIS No. AD/A-
001 072).

Treisman, A. M. Contextual cues in selective listening.
Quarterly Journal of Experimental Psychology, 1960,
_12_, 242-248.

Trumbo, D., & Milone, F. Primary task performance as a
function of encoding, retention, and recall in a
secondary task. _Journal of Experimental Psychology_,
1971, _91_ (2), 273-279.

Trumbo D., Noble, M., & Swink, J. Secondary task inter-
ference in the performance of tracking tasks.
Journal of Experimental Psychology, 1967, _73_ (2),
232-240.

Welford, A. T. _Fundamentals of skill_. London: Methuen
& Co., 1971.

Woodworth, R. S. & Schlosberg, H. _Experimental psychology_
(rev. ed.). New York: Holt, Rhinehart, & Winston,
1954.

APPENDIX: INSTRUCTIONS TO THE SUBJECTS

Introduction

You will be participating in an experiment that will test your performance skills in three types of tasks: 1) You will be asked to use an aircraft-style keyboard to press the same numbered key as the number you will see on the display screen; 2) You will read the number on the screen into a headset microphone that is connected to a voice recognition machine; 3) You will use a video-type game where you are trying to keep a moving target on the right track. Finally, you will be asked to do two tasks at the same time: key-press while you are tracking and verbalize numbers while you are tracking. You will be given detailed instructions before starting each task.

As you were told before you signed up for this experiment, the total time for the tests will be approximately three hours. You are free to get up and move around while the experimenter is setting up the tests in the control room. Since it is important that you keep your voice clear, there will be softdrinks available during the rest periods.

You will be told about your own results at the end of the experiments.

Calibration Trials

Keyboard : This keyboard has four rows of keys with the numbers 1 to 4 in the first row; 5 to 8 in the second row; 9 to 12 in the third row, and 13 to 16 in the last row. You may only use your left hand for the key-pressing. To familiarize yourself with the keyboard press each key in numerical order, and repeat this sequence ten times, so that you know where the different numbers are located. You can press the keys with the tip of your fingers or with the flat of your fingers, whichever is most natural for you. One of the four rows on the keyboard has been designated your "home-row", and you will be told during the practice session to let your fingers go back to the "home-row" after each response. During the actual experiment you will be given other rows as your "home-row".

Voice Recognition Machine : The experimenter will show you how the voice machine works and demonstrate a voice response. Put on the headset with the microphone and start to read the numbers as they appear on the screen. A printout will reveal how well the machine accepted your voice commands. If the machine has misinterpreted a number that you gave, that digit will have to be repeated until the printout shows that the machine recognizes all your numbers correctly.

Practice Trials

Next, there will be five practice trials. The experimenter will read the instructions before the beginning of each type of trial, so that it is totally understood what you are expected to do. During these practice trials you are free to try out any strategy that you feel might work for you, so that you are perfectly comfortable with the equipment. This way you can avoid switching tactics during the actual experiment.

Single-task procedures

Digit processing with manual key-press reponse : In this task you will respond to the number that appears on the screen by pressing the same numbered key. The numbers appear in random order and as soon as you have responded to one number, the next will appear. Even if you make a mistake, go on to the number that you see on the screen. Do not attempt to correct any errors; that will only result in additional errors. Work as quickly and accurately as you can. A bar graph on the left side of the display will show how well you are doing. If you are fast and accurate, the bar increases in height. Try to keep the bar as high as possible.

Are there any questions?

Remember to keep your fingers on the "home-row".

Please start responding as soon as you see the number appear on the screen.

Digit processing with voice response : The same procedures as before, except that you will now name the number you see on the screen instead of pressing the key. It is important that your voice be consistently clear and well projected, so that the machine can recognize and match your voice response. Mumbling, stuttering, and slurring from you will be scored as errors by the machine.

Are there any questions?

Please do not say anything into the microphone until you see the number appear on the screen.

Tracking : This task requires you to keep the diamond-shaped symbol centered within the horizontal track by making appropriate left-right movements of the hand-control Moving the control to the left, moves the diamond to the left; moving the control right, moves the diamond right. There will be random movements of the diamond continuously throughout the trial, and you should respond to these movements as quickly as possible in order to center the diamond on the vertical center line of the track. The bar graph on the right side of the display will indicate how well you are performing. Fast, accurate responding will make the bar increase in height. Try to

keep the bar as high as you can.

Are there any questions?

Please start the task as soon as the diamond is moving.

Dual-task Procedures

Manual Key-press with Tracking : In this part of the experiment you will perform the manual key-pressing for numbers together with the tracking task. The display will now include both tasks, and the two bar graphs will reveal how well you are performing each task. The momentary height represents your average performance over the last few seconds. There will also be a goal box for each task corresponding to your previous performance level, which you should attempt to reach and maintain with the two bar graphs. Note that the height of the goal box is the same for both tasks, indicating that the tasks are of equal importance.

Are there any questions?

Please start as soon as the diamond is moving and the number appears.

Vocal Response with Tracking : The same procedures as for manual key-press with tracking, except that you will now name the number at the same time as you are tracking. Try to keep the two bar graphs as high and as even as possible.

Are there any questions?

Please start as soon as the diamond is moving and the number appears.

Experimental Trials

The practice sessions are now over. During the rest of the experiment you will be performing the same tasks as you have just learned. Before each trial the experimenter will inform you which task or combination of tasks you are to perform.

__Manual__ : Your task this time is the manual key-pressing. You will be presented with the digits ____ through ____. Remember to keep your fingers on the home-row of the numbers ____ through ____. Try to keep the bar as high as you can. Get set. Go.

__Vocal__ : Your task is now to name the numbers. You will be presented with the digits ____ through ____. Remember to speak clearly, but do not say anything until the number appears. Keep the bar graph as high as you can. Get set. Go.

__Tracking__ : You are now going to do the tracking task. Start as soon as the diamond is moving and try to keep the bar as high as you can. Get ready. Go.

__Manual with Tracking__ : You are now going to do key-pressing and tracking at the same time.

digits this time will be ____ through ____. Your home-row is now ____ through ____. Again, try to keep the two bars as even and as high as possible. Please start as soon as the diamond is moving and the number appears. Get ready. Go.

Vocal with Tracking : You are now going to name the numbers and track at the same time. The screen will show the numbers ____ through ____. Do not speak until the number appears and the diamond is moving. Try to keep both bars as even and as high as possible. Get ready. Go.

Debriefing

Subjects were thanked and paid, and given an explanation of the purpose of the study. Individual results were discussed and questions answered.

LME 8